

VI.1 Ιδιοτιμές - Ιδιοδιανύσματα

Αν έχουμε ένα διανυσματικό χώρο V , με μία γραμμική συνάρτηση $L : V \rightarrow V$ και ένα υπόχωρο $W \subset V$, τότε ο περιορισμός της L στον W δεν δίνει πάντα μια συνάρτηση $W \rightarrow W$. Πράγματι, θα μπορούσε να υπάρχει $w \in W$ με $L(w) \in L \setminus W$. Για παράδειγμα η γραμμική συνάρτηση $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ η οποία ως προς την κανονική βάση $= \{e_1, e_2\}$ έχει πίνακα

$$(L, B, B) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

στέλνει το e_1 στο $L(e_1) = e_1 + e_2$, συνεπώς η L αν περιοριστεί στον $W = \langle e_1 \rangle$ δεν δίνει συνάρτηση $W \rightarrow W$.

Ορισμός VI.1.1. Έστω V ένας $\mathbb{F}$ -διανυσματικός χώρος και $L : V \rightarrow V$ μια γραμμική απεικόνιση. Ένας υπόχωρος $W \subset V$ θα λέγεται L -αναλλοίωτος αν και μόνο αν $L(W) \subset W$.

Παράδειγμα VI.1.2. Οι τετριμμένοι υπόχωροι $V \subset V$ και $\{0\} \subset V$ είναι L -αναλλοίωτοι.

Λήμμα VI.1.3. Ένας μονοδιάστατος υπόχωρος $\langle v \rangle \subset V$ είναι L -αναλλοίωτος αν και μόνο αν $Lv = \lambda v$, για κάποιο $\lambda \in \mathbb{F}$.

Απόδειξη. Αν $Lv = \lambda v$ για κάποιο $\lambda \in \mathbb{F}$ τότε $L(\langle v \rangle) \subset \langle v \rangle$. Αντιστρόφως αν $L(\langle v \rangle) \subset \langle v \rangle$, τότε $L(v) \in \langle v \rangle$, άρα $L(v) = \lambda v$ για κάποιο $\lambda \in \mathbb{F}$. $\square$

Ορισμός VI.1.4. Έστω V ένας $\mathbb{F}$ -διανυσματικός χώρος και $L : V \rightarrow V$ μία γραμμική απεικόνιση. Ένα $\lambda \in \mathbb{F}$ θα λέγεται ιδιοτιμή αν υπάρχει μή μηδενικό $v \in V$, ώστε $L(v) = \lambda v$. Το διάνυσμα v θα λέγεται ιδιοδιάνυσμα ως προς την ιδιοτιμή λ . Ο χώρος όλων των ιδιοδιανυσμάτων μιας ιδιοτιμής λ , θα λέγεται ιδιόχωρος της λ . Θα συμβολίζουμε τον ιδιόχωρο της ιδιοτιμής λ με E_λ .

Παρατήρηση VI.1.5. Ένα στοιχείο $v \in E_\lambda$ αν και μόνο αν $Lv - \lambda \text{Id}_V v = 0$ αν και μόνο αν $v \in \text{Ker}(L - \lambda \text{Id}_V)$. Συνεπώς $E_\lambda = \text{Ker}(L - \lambda \text{Id}_V)$ είναι υπόχωρος του V . Θέλουμε ο υπόχωρος αυτός να μην είναι ο μηδενικός και αυτό συμβαίνει αν και μόνο αν $\det(L - \lambda \text{Id}_V) \neq 0$.

Λήμμα VI.1.6. Έστω $L : V \rightarrow V$ γραμμική συνάρτηση, λ_1, λ_2 ιδιοτιμές και $E_{\lambda_1}, E_{\lambda_2}$ οι ιδιόχωροι. Αν $\lambda_1 \neq \lambda_2$ έχουμε $E_{\lambda_1} \cap E_{\lambda_2} = \{0_V\}$.

Απόδειξη. Αν $v \in E_{\lambda_1}$ τότε $Lv = \lambda_1 v$ και αν $v \in E_{\lambda_2}$ τότε $Lv = \lambda_2 v$. Συνεπώς $(\lambda_1 - \lambda_2)v = 0_V$ και αφού $\lambda_1 - \lambda_2 \neq 0$ καταλήγουμε στο ότι $v = 0_V$. $\square$

Ορισμός VI.1.7. Αν ο χώρος V έχει πεπερασμένη διάσταση τότε το πολυώνυμο

$$\text{Ch}_L(x) = \det(L - x\text{Id}_V)$$

θα λέγεται χαρακτηριστικό πολυώνυμο της L .

Παρατήρηση VI.1.8. Αν B είναι μια βάση του V , και $A = [L, B, B]$ είναι ο πίνακας της L ως προς την βάση B , τότε

$$\text{Ch}_L(x) = \det(A - x\mathbb{I}_n). \quad (\text{VI.1})$$

Παρατηρούμε ότι η έκφραση στην εξίσωση (VI.1) είναι ανεξάρτητη της επιλογής της βάσης B . Πράγματι, για μια διαφορετική επιλογή βάσης B' , αντί για τον πίνακα A θα είχαμε τον πίνακα $Q A Q^{-1}$, Q αντιστρέψιμος. Όμως

$$\det(Q A Q^{-1} - x\mathbb{I}_n) = \det(Q A Q^{-1} - x Q \mathbb{I}_n Q^{-1}) = \det(Q) \det(A - x\mathbb{I}_n) \det(Q^{-1}) = \det(A - x\mathbb{I}_n).$$

Η ιδιότητα αυτή εκφράζεται συχνά και ως: «όμοιοι πίνακες έχουν το ίδιο χαρακτηριστικό πολυώνυμο».

Λήμμα VI.1.9. Αν $A \in \mathbb{F}^{n,n}$, η ποσότητα

$$\det(A - x\mathbb{I}_n)$$

είναι ένα πολυώνυμο βαθμού n .

Απόδειξη. Υπολογίζουμε το χαρακτηριστικό πολυώνυμο με την βοήθεια της πρότασης III.6.4

$$\det(A - x\mathbb{I}_n) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)},$$

όπου $b_{ij} = a_{ij} - x\delta_{ij}$. Είναι σαφές ότι ο μεγαλύτερος βαθμός του πολυωνύμου $b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)}$ είναι n , οπότε το ζητούμενο έπεται. $\square$

Παρατήρηση VI.1.10. Ο σταθερός όρος του χαρακτηριστικού πολυωνύμου είναι η ορίζουσα του L , αφού $\text{Ch}_L(0) = \det(L)$.

Παρατήρηση VI.1.11. Οι ρίζες λ του χαρακτηριστικού πολυωνύμου είναι ακριβώς οι τιμές $\lambda \in \mathbb{F}$, για τις οποίες το ομογενές σύστημα

$$L v - \lambda v = 0$$

έχει, εκτός από την μηδενική λύση και άλλες λύσεις, δηλαδή οι ιδιοτιμές.

Λήμμα VI.1.12. Έστω $\lambda_1, \dots, \lambda_n$ ανά δύο διαφορετικές ιδιοτιμές μιας γραμμικής συνάρτησης $L : V \rightarrow V$ με αντίστοιχα ιδιοδιανύσματα $v_1, \dots, v_n$. Τότε τα $v_1, \dots, v_n$ είναι γραμμικά ανεξάρτητα.

Απόδειξη. Θα δόσουμε μια επαγωγική απόδειξη στο πλήθος n των ιδιοδιανυσμάτων $v_1, \dots, v_n$. Αν $n = 1$, τότε το $v_1 \neq 0$ είναι γραμμικά ανεξάρτητο. Υποθέτουμε ότι ισχυρισμός είναι αληθής για $n - 1$ το πλήθος ιδιοδιανύσματα. Έστω $v_1, \dots, v_n$ ιδιοδιανύσματα τα οποία αντιστοιχούν στις διαφορετικές ανά δύο ιδιοτιμές $\lambda_1, \dots, \lambda_n$. Έστω ότι αυτά ικανοποιούν μια σχέση

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n = 0. \quad (\text{VI.2})$$

Εφαρμόζουμε την L για να πάρουμε

$$a_1 L(v_1) + a_2 L(v_2) + \dots + a_n L(v_n) = a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \dots + a_n \lambda_n v_n = L(0_V) = 0_V. \quad (\text{VI.3})$$

Πολλαπλασιάζουμε την σχέση (VI.2) με λ_n και αφαιρούμε την σχέση (VI.3) για να πάρουμε:

$$a_1 (\lambda_n - \lambda_1) v_1 + \dots + (\lambda_n - \lambda_{n-1}) v_{n-1} = 0.$$

Από την επαγωγική υπόθεση έχουμε ότι τα $v_1, \dots, v_{n-1}$ είναι γραμμικά ανεξάρτητα, συνεπώς $a_v (\lambda_n - \lambda_v) = 0$ για κάθε $1 \leq v \leq n - 1$. Επειδή $\lambda_n - \lambda_v \neq 0$ θα πρέπει $a_1 = \dots = a_{n-1} = 0$. Τότε από την σχέση (VI.2) έχουμε ότι και $a_n = 0$, άρα τα $v_1, \dots, v_n$ είναι γραμμικά ανεξάρτητα. $\square$

Πόρισμα VI.1.13. Έστω ότι το χαρακτηριστικό πολυώνυμο $\text{Ch}_L(x)$ έχει n -το πλήθος διαφορετικές ρίζες, οι οποίες είναι ιδιοτιμές. Τότε υπάρχει μια βάση $B = \{v_1, \dots, v_n\}$ αποτελούμενη από ιδιοδιανύσματα ώστε

$$(L, B, B) = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

Απόδειξη. Τα n το πλήθος ιδιοδιανύσματα $v_1, \dots, v_n$ που αντιστοιχούν στις διαφορετικές ανά δύο ιδιοτιμές είναι γραμμικά ανεξάρτητα και είναι τόσα όσα η διάσταση, άρα αποτελούν μια βάση B του V , σύμφωνα με το θεώρημα IV.3.11. Η διαγώνια μορφή του πίνακα (L, B, B) προκύπτει από το ότι $Lv_i = \lambda_i v_i$ για κάθε $1 \leq i \leq n$. $\square$

Παρατήρηση VI.1.14. Αν δουλέψουμε σε ένα αλγεβρικά κλειστό σώμα τότε το χαρακτηριστικό πολυώνυμο της $L : V \rightarrow V$ έχει $\dim V = n$ το πλήθος ρίζες.

Ορισμός VI.1.15. Ένας πίνακας $A \in \mathbb{F}^{n,n}$ λέγεται διαγωνοποιήσιμος αν υπάρχει αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{n,n}$ ώστε

$$Q^{-1} A Q = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Παρατήρηση VI.1.16. Υπάρχουν πίνακες που δεν είναι διαγωνοποιήσιμοι. Για παράδειγμα ο πίνακας

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

δεν είναι διαγωνοποιήσιμος. Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του A είναι $(x-1)^2$. Πράγματι αν υπήρχε αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{2,2}$ ώστε

$$Q^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} Q = \text{diag}(\lambda_1, \lambda_2), \quad (\text{VI.4})$$

τότε $\text{Ch}_{\text{diag}(\lambda_1, \lambda_2)}(x) = (x - \lambda_1)(x - \lambda_2)$ το οποίο όμως θα πρέπει να είναι ίσο με το $(x - 1)^2$ αφού όμοιοι πίνακες έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Συνεπώς $\lambda_1 = \lambda_2 = 1$ και η εξίσωση (VI.4) γίνεται

$$Q^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} Q = \mathbb{I}_2 \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \mathbb{I}_2,$$

το οποίο είναι άτοπο.

Ορισμός VI.1.17. Θεωρούμε μια ιδιοτιμή λ της γραμμικής συνάρτησης $L : V \rightarrow V$.

- Θα ονομάζουμε *αλγεβρική πολλαπλότητα* της ιδιοτιμής λ τον φυσικό αριθμό $\mu(\lambda)$ για τον οποίο ισχύει

$$\text{Ch}_L(x) = (x - \lambda)^{\mu(\lambda)} q(x),$$

όπου $q(x) \in \mathbb{F}[x]$ με $q(\lambda) \neq 0$, δηλαδή την πολλαπλότητα της ρίζας λ στο χαρακτηριστικό πολυώνυμο.

- Θα ονομάζουμε *γεωμετρική πολλαπλότητα* $\rho(\lambda)$ της ιδιοτιμής λ την διάσταση $\dim E_\lambda$ του ιδιόχωρου της ιδιοτιμής λ .

Πρόταση VI.1.18. Για μια ιδιοτιμή λ της γραμμικής συνάρτησης L η αλγεβρική πολλαπλότητα $\mu(\lambda)$ και η γεωμετρική πολλαπλότητα $\rho(\lambda)$ ικανοποιούν την σχέση

$$1 \leq \rho(\lambda) \leq \mu(\lambda) \leq \dim(V) = n.$$

Απόδειξη. Η πρώτη και η τελευταία ανισότητα είναι προφανής. Θα αποδείξουμε ότι $\rho(\lambda) \leq \mu(\lambda)$. Θεωρούμε μια βάση $\{v_1, \dots, v_{\rho(\lambda)}\}$ του ιδιόχωρου E_λ την οποία επεκτείνουμε σε μια βάση $B = \{v_1, \dots, v_{\rho(\lambda)}, w_1, \dots, w_k\}$ του V . Γράφουμε τον πίνακα της L ως προς την βάση B

$$(L, B, B) = \begin{pmatrix} \lambda & 0 & \cdots & 0 & a_{1, \rho(\lambda)+1} & \cdots & a_{1n} \\ 0 & \lambda & \ddots & \vdots & a_{2, \rho(\lambda)+1} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda & a_{\rho(\lambda), \rho(\lambda)+1} & \cdots & a_{\rho(\lambda), n} \\ 0 & \cdots & \cdots & 0 & a_{\rho(\lambda)+1, \rho(\lambda)+1} & \cdots & a_{\rho(\lambda)+1, n} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{n, \rho(\lambda)+1} & \cdots & a_{n, n} \end{pmatrix} = \left(\begin{array}{c|c} \lambda \mathbb{I}_{\rho(\lambda)} & A \\ \hline \mathbf{0}_{n-\rho(\lambda), \rho(\lambda)} & A' \end{array} \right).$$

Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του παραπάνω πίνακα

$$\det \left(\begin{array}{c|c} (\lambda - x) \mathbb{I}_{\rho(\lambda)} & A \\ \hline \mathbf{0}_{n-\rho(\lambda), \rho(\lambda)} & A' - x \mathbb{I}_{n-\rho(\lambda)} \end{array} \right) = (x - \lambda)^{\rho(\lambda)} (-1)^{\rho(\lambda)} g(x).$$

διαίρεται από το $(x - \lambda)^{\rho(\lambda)}$, από όπου προκύπτει η ζητούμενη ανισότητα, η οποία μπορεί να είναι γνήσια αν $g(\lambda) = 0$. $\square$

Πρόταση VI.1.19. Ένας πίνακας είναι διαγωνοποιήσιμος αν και μόνο αν για κάθε ιδιοτιμή $\rho(\lambda) = \mu(\lambda)$.

Απόδειξη. Ας υποθέσουμε ότι πίνακας $A \in \mathbb{F}^{n,n}$ είναι διαγωνοποιήσιμος. Υπάρχει μια βάση $v_1, \dots, v_n$ ώστε $Av_i = \lambda_i v_i$, δηλαδή μια βάση αποτελούμενη από ιδιοδιανύσματα. Έστω $\lambda_{i_1}, \dots, \lambda_{i_r}$ οι ανά δύο διαφορετικές ιδιοτιμές. Το χαρακτηριστικό πολυώνυμο υπολογίζεται ως

$$\text{Ch}_A(x) = \prod_{v=1}^s (\lambda_{i_v} - x)^{\mu(\lambda_{i_v})}.$$

Είναι σαφές ότι

$$n = \dim \mathbb{F}^n = \sum_{v=1}^s \rho(\lambda_{i_v}) \leq \sum_{v=1}^s \mu(\lambda_{i_v}) = \deg \text{Ch}_A(x) = n$$

από όπου προκύπτει ότι $\rho(\lambda) = \mu(\lambda)$ για κάθε ιδιοτιμή λ .

Αντιστρόφως αν για κάθε ιδιοτιμή λ έχουμε $\rho(\lambda) = \mu(\lambda)$ τότε μπορούμε να βρούμε

$$n = \deg \text{Ch}_A(x) = \sum_{v=1}^s \mu(\lambda_{i_v}) = \sum_{v=1}^s \rho(\lambda_{i_v})$$

το πλήθος γραμμικά ανεξάρτητα ιδιοδιανύσματα τα οποία αποτελούν μια βάση του χώρου $\mathbb{F}^n$ και συνεπώς ο πίνακας A είναι διαγωνοποιήσιμος. $\square$

Παρατήρηση VI.1.20. Υποθέτουμε ότι ο πίνακας $A \in \mathbb{F}^{n,n}$ είναι διαγωνοποιήσιμος, δηλαδή υπάρχει ένας αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{n,n}$ και διαγώνιος πίνακας $\Delta = \text{diag}(\lambda_1, \dots, \lambda_n)$ ώστε

$$Q^{-1}AQ = \Delta \Leftrightarrow AQ = Q\Delta.$$

Αν γράψουμε $Q = (q^1, q^2, \dots, q^n)$, όπου $q^1, \dots, q^n$ είναι οι γραμμικά ανεξάρτητες στήλες του Q , τότε έχουμε $Aq^v = \lambda_v q^v$, για $1 \leq v \leq n$, δηλαδή ο πίνακας Q έχει ως στήλες ιδιοδιανύσματα.

Μέθοδος 3 (Διαγωνοποίηση). Για να διαγωνοποιήσουμε ένα $n \times n$ πίνακα εκτελούμε τα παρακάτω βήματα:

- 1 Υπολογισμός χαρακτηριστικού πολυωνύμου
- 2 Εύρεση ιδιοτιμών $\lambda_1, \dots, \lambda_n$
- 3 Υπολογισμός βάσεων B_i των ιδιοχώρων E_{λ_i} .
- 4 Ο πίνακας Q είναι ένας πίνακας που έχει ως στήλες τα στοιχεία των βάσεων B_i .

Παράδειγμα VI.1.21. Θέλουμε να διαγωνοποιήσουμε τον πίνακα

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\text{Ch}_A(x) = \det(A - x\mathbb{I}_2) = \det \begin{pmatrix} -x & 1 \\ 1 & 1-x \end{pmatrix} = x^2 - x - 1 = \left(x - \frac{1-\sqrt{5}}{2}\right) \left(x + \frac{1+\sqrt{5}}{2}\right)$$

Συνεπώς οι ιδιοτιμές είναι οι $\frac{1-\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2}$. Στην συνέχεια υπολογίζουμε τους ιδιοχώρους. Για την ιδιοτιμή $\frac{1-\sqrt{5}}{2}$ ο $E_{\frac{1-\sqrt{5}}{2}}$ είναι ο χώρος λύσεων της

$$\left(A - \frac{1-\sqrt{5}}{2}\mathbb{I}_2\right)x = 0 \text{ ισοδύναμα } \begin{pmatrix} \frac{1}{2}(\sqrt{5}-1) & 1 \\ 1 & 1+\frac{1}{2}(\sqrt{5}-1) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

το οποίο λύνουμε με την μέθοδο απαλοιφής του Gauss και παίρνουμε ως βάση το χώρου λύσεων το $((-1 - \sqrt{5})/2, 1)^t$.

Ομοίως για την ιδιοτιμή $\frac{1+\sqrt{5}}{2}$ ο $E_{\frac{1+\sqrt{5}}{2}}$ είναι ο χώρος λύσεων της

$$\left(A - \frac{1+\sqrt{5}}{2}\mathbb{I}_2\right)x = 0 \text{ ισοδύναμα } \begin{pmatrix} \frac{1}{2}(-1-\sqrt{5}) & 1 \\ 1 & 1 + \frac{1}{2}(-1-\sqrt{5}) \end{pmatrix}$$

ο οποίος έχει βάση του χώρου λύσεων το $((-1 + \sqrt{5})/2, 1)^t$

Σχηματίζουμε τον πίνακα Q ο οποίος έχει ως πρώτη στήλη το μοναδικό στοιχείο της βάσης του ιδιόχωρου της ιδιοτιμής $(1 - \sqrt{5})/2$ και ως δεύτερη στήλη το μοναδικό στοιχείο της βάσης του ιδιόχωρου της ιδιοτιμής $(1 + \sqrt{5})/2$, δηλαδή

$$Q = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & \frac{-1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$$

Υπολογίζουμε ότι

$$Q^{-1} = \begin{pmatrix} -\frac{1}{\sqrt{5}} & -\frac{1-\sqrt{5}}{2\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{-1-\sqrt{5}}{2\sqrt{5}} \end{pmatrix}$$

και τέλος επαληθεύουμε ότι

$$Q^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} Q = \begin{pmatrix} \frac{(1-\sqrt{5})}{2} & 0 \\ 0 & \frac{(1+\sqrt{5})}{2} \end{pmatrix}. \quad (\text{VI.5})$$

Παράδειγμα VI.1.22. Ας προσπαθήσουμε να διαγωνοποιήσουμε τον πίνακα με την παραπάνω μέθοδο.

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\text{Ch}_A(x) = \det(A - x\mathbb{I}_3) = -(x-2)^3.$$

Η μοναδική ιδιοτιμή είναι το 2. Στην συνέχεια υπολογίζουμε τον ιδιόχωρο E_2 ο οποίος δίνεται από τις λύσεις του ομογενούς συστήματος

$$(A - 2\mathbb{I}_3)x = 0 \text{ ισοδύναμα } \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

το οποίο λύνουμε με την μέθοδο απαλοιφής του Gauss, ο πίνακας είναι ήδη σε κλιμακωτή μορφή και ο χώρος λύσεων είναι μονοδιάστατος με βάση $(1, 0, 0)^t$. Δεν μπορούμε να βρούμε μια βάση του $\mathbb{R}^3$ αποτελούμενη από ιδιοδιανύσματα, οπότε ο πίνακας A , δεν είναι διαγωνοποιήσιμος.

Παράδειγμα VI.1.23. Θέλουμε να διαγωνοποιήσουμε τον πίνακα

$$A = \begin{pmatrix} -3 & 6 & -2 & 6 \\ 16 & -49 & 22 & -46 \\ 7 & -30 & 15 & -27 \\ -15 & 42 & -18 & 40 \end{pmatrix}$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\text{Ch}_A(x) = \det(A - x\mathbb{I}_4) = -2 + 3x + x^2 - 3x^3 + x^4$$

Αυτός γενικά είναι ένας δύσκολος και απαιτητικός υπολογισμός αν ο πίνακας A δεν είναι κάποιας ειδικής μορφής.

Στην συνέχεια υπολογίζουμε τις ιδιοτιμές. Αυτό για ένα πίνακα $n \times n$ προϋποθέτει ότι μπορούμε να υπολογίσουμε τις ρίζες ενός πολυωνύμου n -στου βαθμού. Μια τέτοια λύση εν γένει δεν μπορεί να γίνει με ακριβή τρόπο αν το $n \geq 5$ και στην πράξη χρησιμοποιούμε μεθόδους αριθμητικής ανάλυσης για να υπολογίσουμε ρητές προσεγγίσεις των ριζών. Στο συγκεκριμένο παράδειγμα έχουμε ότι

$$\text{Ch}_A(x) = (x-2)(x-1)^2(x+1).$$

Οι ιδιοτιμές είναι οι 2, 1, -1 και το χαρακτηριστικό πολυώνυμο έχει την ρίζα -1 με πολλαπλότητα 2, οπότε δεν μπορούμε να χρησιμοποιήσουμε το πόρισμα VI.1.13 για να εξασφαλίσουμε ότι ο πίνακας A είναι διαγωνοποιήσιμος. Ο πίνακας A μπορεί να είναι διαγωνοποιήσιμος μπορεί και όχι και πρέπει να το ερευνήσουμε περισσότερο.

Υπολογίζουμε βάσεις για τους ιδιόχωρους των ιδιοτιμών:

Για την ιδιοτιμή 2 Ο ιδιόχωρος E_2 είναι η λύση του ομογενούς συστήματος

$$(A - 2I_4)x = 0 \text{ ισοδύναμα } \begin{pmatrix} -5 & 6 & -2 & 6 \\ 16 & -51 & 22 & -46 \\ 7 & -30 & 13 & -27 \\ -15 & 42 & -18 & 38 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Το παραπάνω σύστημα το λύνουμε με την μέθοδο της απαλοιφής του Gauss για να βρούμε μια βάση του ιδιόχωρου E_2 η οποία δίνεται από το διάνυσμα $(0, -4, -3, 3)^t$

Για την ιδιοτιμή 1 Ο ιδιόχωρος E_1 είναι η λύση του ομογενούς συστήματος

$$(A - I_4)x = 0 \text{ ισοδύναμα } \begin{pmatrix} -4 & 6 & -2 & 6 \\ 16 & -50 & 22 & -46 \\ 7 & -30 & 14 & -27 \\ -15 & 42 & -18 & 39 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Και αυτό το σύστημα το λύνουμε με την μέθοδο της απαλοιφής του Gauss για να βρούμε μια βάση του ιδιόχωρου E_1 η οποία δίνεται από τα διανύσματα $(3, -11, 0, 13)^t, (4, 7, 13, 0)^t$.

Για την ιδιοτιμή -1 Ο ιδιόχωρος E_{-1} είναι η λύση του ομογενούς συστήματος

$$(A + I_4)x = 0 \text{ ισοδύναμα } \begin{pmatrix} -2 & 6 & -2 & 6 \\ 16 & -48 & 22 & -46 \\ 7 & -30 & 16 & -27 \\ -15 & 42 & -18 & 41 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Το παραπάνω σύστημα το λύνουμε με την μέθοδο της απαλοιφής του Gauss για να βρούμε μια βάση του ιδιόχωρου E_{-1} η οποία δίνεται από το διάνυσμα $(1, -3, -1, 3)^t$.

Σχηματίζουμε τον πίνακα

$$Q = \begin{pmatrix} 0 & 3 & 4 & 1 \\ -4 & -11 & 7 & -3 \\ -3 & 0 & 13 & -1 \\ 3 & 13 & 0 & 3 \end{pmatrix}$$

Η πρώτη στήλη είναι το ιδιοδιάνυσμα της ιδιοτιμής 2, οι δεύτερη και η τρίτη στήλη είναι βάση του ιδιόχωρου της ιδιοτιμής 1 και η τέταρτη στήλη είναι ιδιοδιάνυσμα της ιδιοτιμής -1.

Μπορούμε τώρα να επαληθεύσουμε ότι

$$Q^{-1} = \begin{pmatrix} -1 & 8 & -4 & 7 \\ -\frac{3}{13} & -\frac{15}{13} & \frac{9}{13} & -\frac{11}{13} \\ -\frac{1}{13} & \frac{21}{13} & -\frac{10}{13} & \frac{18}{13} \\ 2 & -3 & 1 & -3 \end{pmatrix}$$

και επαληθεύουμε ότι

$$Q^{-1}AQ = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Ασκήσεις

Άσκηση VI.1 Ναδειχθεί ότι κάθε άνω τριγωνικός πίνακας $A = (a_{ij})$ έχει χαρακτηριστικό πολυώνυμο το

$$\text{Ch}_A(x) = \prod_{i=1}^n (a_{ii} - x).$$

Άσκηση VI.2 Να αποδειχθεί ότι οι πίνακες $A, A^t \in \mathbb{F}^{n,n}$ έχουν το ίδιο χαρακτηριστικό πολυώνυμο.

Άσκηση VI.3 Δίνεται ο πίνακας $A = (a_{ij}) \in \mathbb{F}^{n,n}$ με $a_{ij} = 1$ για κάθε $1 \leq i, j \leq n$. Να βρεθεί αντιστρέψιμος πίνακας Q ώστε $Q^{-1}AQ = \Delta$, όπου Δ διαγώνιος πίνακας.

Άσκηση VI.4 Θεωρούμε τον πίνακα

$$A = \begin{pmatrix} a & b & \cdots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{pmatrix}$$

Να βρεθεί αντιστρέψιμος πίνακας Q ώστε $Q^{-1}AQ = \Delta$, όπου Δ διαγώνιος πίνακας.

Άσκηση VI.5 Να υπολογιστεί το χαρακτηριστικό πολυώνυμο του πίνακα

$$A = \begin{pmatrix} 0 & a & \cdots & a \\ b & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ b & \cdots & b & 0 \end{pmatrix}.$$

Άσκηση VI.6 Δίνονται οι πίνακες $A \in \mathbb{F}^{m,n}$, $B \in \mathbb{F}^{n,m}$. Αποδείξτε ότι

$$x^n(-1)^n \text{Ch}_{AB}(x) = x^n \det(x\mathbb{I}_m - AB) = x^m \det(x\mathbb{I}_m - BA) = x^m(-1)^m \text{Ch}_{BA}(x)$$

Άσκηση VI.7 Στην άσκηση [VI.6](#) αν $n = m$ ναδειχθεί ότι οι πίνακες AB και BA έχουν το ίδιο χαρακτηριστικό πολυώνυμο.

Άσκηση VI.8 Να δείξετε ότι η γραμμική συνάρτηση $T : V \rightarrow V$, όπου V είναι ο διανυσματικός χώρος των συνεχών συναρτήσεων $f : \mathbb{R} \rightarrow \mathbb{R}$ και δίνεται από $Tf(x) = \int_0^x f(t)dt$ δεν έχει ιδιοτιμές.

Άσκηση VI.9 Θεωρούμε την γραμμική συνάρτηση $D : V \rightarrow V$, όπου V είναι ο διανυσματικός χώρος των διαφορίσιμων συναρτήσεων $f : \mathbb{R} \rightarrow \mathbb{R}$ και $Df = f'$. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα της f .

Άσκηση VI.10 Να αποδειχθεί ότι $(2 + \sqrt{3})^n + (2 - \sqrt{3})^n \in \mathbb{Z}$ για κάθε $n \in \mathbb{N}$.

Άσκηση VI.11 Δείξτε ότι αν ο πίνακας A είναι αντιστρέψιμος και λ είναι μια ιδιοτιμή του A τότε λ^{-1} είναι μια ιδιοτιμή του A^{-1} .

Άσκηση VI.12 Αν ο πίνακας $A \in \mathbb{F}^{n,n}$ είναι αντιστρέψιμος, είναι όμοιος με τον A^{-1} και ο n είναι περιττός τότε έχει μια ιδιοτιμή $\lambda = \pm 1$.

Άσκηση VI.13 Δείξτε ότι κάθε πίνακας A με τάξη 1 έχει χαρακτηριστικό πολυώνυμο

$$\text{Ch}_A(x) = (-x)^{n-1}(\text{tr}(A) - x)$$

Στην συνέχεια υπολογίστε το χαρακτηριστικό πολυώνυμο του πίνακα

$$A = \begin{pmatrix} a & b & \cdots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{pmatrix}$$

VI.2 Το θεώρημα Caley-Hamilton

Ορισμός VI.2.1. Έστω $f(x) \in \mathbb{F}[x]$,

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n.$$

για ένα τετραγωνικό πίνακα $A \in \mathbb{F}^{m,m}$ μπορούμε να θεωρήσουμε τον πίνακα

$$f(A) = a_0\mathbb{I}_m + a_1A + a_2A^2 + \cdots + a_nA^n.$$

Παρατηρούμε ότι ο διανυσματικός χώρος $\mathbb{F}^{m,m}$ έχει διάσταση m^2 , οπότε οποιαδήποτε $m^2 + 1$ στοιχεία του είναι γραμμικά εξαρτημένα. Ειδικότερα, τα $m^2 + 1$ στοιχεία

$$\mathbb{I}_m, A, A^2, A^3, \dots, A^{m^2}$$

είναι γραμμικά εξαρτημένα, δηλαδή υπάρχουν στοιχεία $\lambda_0, \lambda_1, \dots, \lambda_{m^2+1} \in \mathbb{F}$ όχι όλα 0 ώστε

$$\lambda_0\mathbb{I}_m + \lambda_1A + \dots + \lambda_{m^2+1}A^{m^2} = \mathbf{0}_{\mathbb{F}^{m,m}}.$$

Άρα για το μη-μηδενικό πολυώνυμο $f(x) = \sum_{v=0}^{m^2} \lambda_v x^v$ έχουμε ότι $f(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$. Θεωρούμε το σύνολο των πολυωνύμων τα οποία μηδενίζουν τον πίνακα A .

$$I_A = \{f \in \mathbb{F}[x] : f(A) = \mathbf{0}_{\mathbb{F}^{m,m}}\}$$

Πρόταση VI.2.2. Το σύνολο I_A είναι ένα ιδεώδες του δακτυλίου $\mathbb{F}[x]$, δηλαδή

1. Περιέχει μη-μηδενικά πολυώνυμα,
2. Αν $f, g \in I_A$, τότε $f + g \in I_A$

3. Αν $f \in I_A$, $g \in \mathbb{F}[x]$, τότε $fg \in I_A$.

Απόδειξη. Το ότι το σύνολο I_A περιέχει και μη-μηδενικά πολυώνυμα το έχουμε ήδη δείξει. Είναι σαφές ότι αν $f, g \in I_A$ τότε $f(A) = g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$ οπότε και $(f+g)(A) = f(A) + g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$. Ομοίως αν $f \in I_A$ και $g \in \mathbb{F}[x]$, τότε $(fg)(A) = f(A)g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$. $\square$

Πόρισμα VI.2.3. Υπάρχει ένα μονικό πολυώνυμο $p_A(x)$, το οποίο διαιρεί κάθε πολυώνυμο που μηδενίζει τον A . Το πολυώνυμο αυτό θα το λήμε το ελάχιστο πολυώνυμο του πίνακα A .

Απόδειξη. Γνωρίζουμε ότι τα ιδεώδη του $\mathbb{F}[x]$ είναι κύρια (δείτε το θεώρημα I.8.3, οπότε υπάρχει πολυώνυμο $p_A(x)$, ώστε $p_A(x)\mathbb{F}[x] = I_A$. $\square$

Παρατήρηση VI.2.4. Όμοιοι πίνακες έχουν το ίδιο ελάχιστο πολυώνυμο. Πράγματι, αφού $QA^kQ^{-1} = (QAQ^{-1})^k$ παρατηρούμε ότι $f(QAQ^{-1}) = Qf(A)Q^{-1}$, συνεπώς $I_A = I_{QAQ^{-1}}$.

Παραδείγματα VI.2.5. • Θεωρούμε ένα διαγωνοποιήσιμο πίνακα, ο οποίος εξ ορισμού είναι όμοιος με ένα διαγώνιο. Οπότε σύμφωνα με την παρατήρηση VI.2.4 αρκεί να υπολογίσουμε το ελάχιστο πολυώνυμο του $\text{diag}(\lambda_1, \dots, \lambda_s)$. Παρατηρούμε ότι για κάθε πολυώνυμο $f(x) \in \mathbb{F}[x]$ έχουμε

$$f(\text{diag}(\lambda_1, \dots, \lambda_n)) = \text{diag}(f(\lambda_1), \dots, f(\lambda_r)),$$

οπότε το ελάχιστο πολυώνυμο είναι το

$$p_{\text{diag}(\lambda_1, \dots, \lambda_n)} = \prod_{v=1}^s (x - \lambda_{i_v}),$$

όπου $\{\lambda_{i_1}, \dots, \lambda_{i_s}\}$ είναι το σύνολο των διαφορετικών $\lambda_1, \dots, \lambda_n$.

- Ένας πίνακας θα λέγεται μηδενοδύναμος βαθμού k αν το ελάχιστο πολυώνυμο του είναι το x^k για κάποιο φυσικό αριθμό $k \geq 1$. Για παράδειγμα ο $n \times n$ πίνακας

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

έχει ελάχιστο πολυώνυμο x^n . Πράγματι μπορούμε να γράψουμε τον πίνακα $N = (n_{ij})$ με $n_{ij} = \delta_{i,j+1}$. Για κάθε $0 < r \leq n$ συμβολίζουμε τον πίνακα $N^r = (n_{ij}^{(r)})$ ο οποίος σύμφωνα με την άσκηση II.2 είναι

$$n_{ij}^{(r)} = \sum_{v_1=1}^{\ell_1} \sum_{v_2=1}^{\ell_2} \dots \sum_{v_r=1}^{\ell_r} n_{i, \ell_1} n_{\ell_1, \ell_2} \dots n_{\ell_{r-1}, j},$$

το οποίο είναι μη μηδενικό αν και μόνο αν $\ell_1 = i+1, \ell_2 = \ell_1+1 = i+2, \dots, j = \ell_{r-1}+1 = i+r$. Συνεπώς $n_{ij}^{(r)} = \delta_{i, i+r}$. Άρα για $r = n-1$ έχουμε ότι $n_{1, n-1}^{(n-1)} = 1$, οπότε $N^{n-1} \neq \mathbf{0}_{\mathbb{F}^{n,n}}$. Ενώ $N^n = \mathbf{0}_{\mathbb{F}^{n,n}}$. Μπορούμε να σκεφτόμαστε τον πίνακα N ως ένα πίνακα με μονάδες ακριβώς στην διαγώνιο που είναι μια θέση πάνω από την κύρια διαγώνιο, ενώ σε κάθε δύναμη N^k η μη μηδενική διαγώνιος με τις μονάδες μετακινείται μια θέση προς τα πάνω.

Θεώρημα VI.2.6 (Caley-Hamilton). Το χαρακτηριστικό πολυώνυμο $\text{Ch}_A(x)$ ενός $n \times n$ πίνακα μηδενίζει τον πίνακα A , δηλαδή $\text{Ch}_A(A) = \mathbf{0}_{\mathbb{F}^{n,n}}$.

Απόδειξη. Υποθέτουμε ότι

$$\text{Ch}_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

Παρατηρούμε ότι ισχύει

$$(A - x\mathbb{I}_n)\text{adj}(A - x\mathbb{I}_n) = \det(A - x\mathbb{I}_n)\mathbb{I}_n.$$

Ο πίνακας $\text{adj}(A - x\mathbb{I}_n)$ έχει ως στοιχεία πολυώνυμα βαθμού το πολύ $n - 1$ συνεπώς μπορεί να γραφεί ως:

$$\text{adj}(A - x\mathbb{I}_n) = B_0 + B_1x + \cdots + B_{n-1}x^{n-1}, \text{ όπου } B_0, B_1, \dots, B_{n-1} \in \mathbb{F}^{n,n}.$$

Συνεπώς έχουμε

$$(A - x\mathbb{I}_n)(B_0 + B_1x + \cdots + B_{n-1}x^{n-1}) = (x^n + a_{n-1}x^{n-1} + \cdots + a_0)\mathbb{I}_n.$$

Συγκρίνοντας τους συντελεστές έχουμε

$$\begin{aligned} AB_0 &= a_0\mathbb{I}_n \\ -B_0 + AB_1 &= a_1\mathbb{I}_n \\ &\vdots \\ -B_{n-2} + AB_{n-1} &= a_{n-1}\mathbb{I}_n \\ -B_{n-1} &= \mathbb{I}_n \end{aligned}$$

Πολλαπλασιάζουμε τις παραπάνω σχέσεις με $\mathbb{I}_n, A, \dots, A^n$ και προσθέτουμε κατά μέλη για να πάρουμε

$$0 = a_0\mathbb{I}_n + a_1A + \cdots + a_{n-1}A^{n-1} + A^n,$$

δηλαδή το ζητούμενο. □

Πόρισμα VI.2.7. Το ελάχιστο πολυώνυμο $p_A(x)$ ενός πίνακα A διαιρεί το χαρακτηριστικό πολυώνυμο $\text{Ch}_A(x)$.

Πόρισμα VI.2.8. Κάθε αντιστρέψιμος πίνακας μπορεί να γραφεί ως μια πολυωνυμική έκφραση βαθμού $n - 1$ στον A .

Απόδειξη. Έχουμε ότι

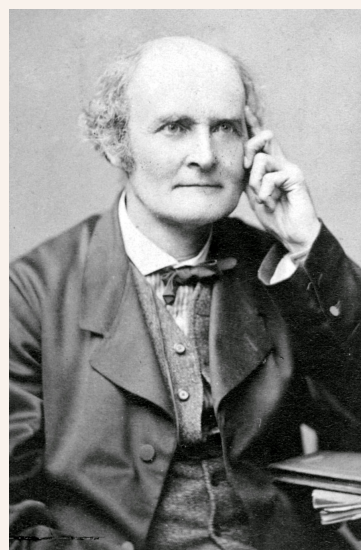
$$\text{Ch}_A(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + (-1)^n \det(A)\mathbb{I}_n = 0$$

άρα

$$A^{-1} = \frac{(-1)^{n+1}}{\det(A)} (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1\mathbb{I}_n).$$

□

Ο **Arthur Cayley** (16 Αυγούστου 1821 - 26 Ιανουαρίου 1895) ήταν Άγγλος πολυγραφώτατος Μαθηματικός. Σπούδασε στο King's College, και στο Trinity College στο Πανεπιστήμιο του Cambridge. Εργάστηκε ως δικηγόρος για 14 έτη. Σε ηλικία 42 ετών έγινε καθηγητής στο Πανεπιστήμιο του Cambridge. Εργάστηκε κυρίως στην Άλγεβρα και ήταν ένας από τους πρώτους Μαθηματικούς που όρισαν την έννοια της ομάδας με τον σύγχρονο τρόπο, ως ένα σύνολο εφοδιασμένο με μια πράξη που ικανοποιεί μια σειρά από ιδιότητες. Οι Μαθηματικοί της εποχής του όταν αναφέρονταν στην έννοια της ομάδας εννοούσαν ομάδες μεταθέσεων. Διάφορα θεωρήματα στην θεωρία ομάδων φέρουν το όνομα του όπως για παράδειγμα το θεώρημα ότι κάθε πεπερασμένη ομάδα είναι υποομάδα μιας ομάδας μεταθέσεων. Ασχολήθηκε και με την Αλγεβρική γεωμετρία, την θεωρία των επιφανειών, τις πολλαπλότητες Chow, τις ruled surfaces αλλά και με την θεωρία των ελλειπτικών, αβελιανών και θήτα συναρτήσεων.



Ο **William Rowan Hamilton** (3/4 Αυγούστου 1805 - 2 Σεπτεμβρίου 1865) ήταν ένας Ιρλανδός Μαθηματικός, Αστρονόμος και Φυσικός. Έδειξε το ταλέντο του από μικρή ηλικία τόσο στα Μαθηματικά όσο και στην εκμάθηση ξένων γλωσσών. Σπούδασε στο κολέγιο Trinity του Δουβλίνου. Στην ηλικία των 21 ετών και ενώ ήταν ακόμα προπτυχιακός φοιτητής του δόθηκε θέση καθηγητή Αστρονομίας στο Trinity college που φοιτούσε. Εργάστηκε σε θέματα Άλγεβρας και Γραμμικής Άλγεβρας, εισήγαγε την έννοια των quaternions και επίσης μελέτησε την πεμπτοβάθμια εξίσωση και τα νέα αποτελέσματα του Niels Henrik Abel. Επίσης είχε μέγιστη συνεισφορά στην οπτική και στην κλασική μηχανική εισάγοντας την έννοια της Χαμιλτονιανής μια θεμελιώδη έννοια της θεωρητικής φυσικής με εφαρμογές στον ηλεκτρομαγνητισμό και στην κβαντομηχανική.



VI.2.1 Εφαρμογές

Πρόταση VI.2.9. Για ένα πίνακα A θεωρούμε το χαρακτηριστικό του πολυώνυμο

$$\text{Ch}_A(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$$

Παρατηρούμε ότι αν $z_1, \dots, z_n$ είναι οι ιδιοτιμές του A τότε οι εκφράσεις

$$z_1^k + \dots + z_n^k = \text{tr}(A^k).$$

VI.2.2 Συνοδός πίνακας

Ορισμός VI.2.10. Για ένα πολυώνυμο

$$f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n \in \mathbb{F}[x]$$

ο παρακάτω $n \times n$ πίνακας

$$C(f) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

θα λέγεται συνοδός πίνακας (companion matrix) του $f(x)$.

Πρόταση VI.2.11. Το χαρακτηριστικό πολυώνυμο του $C(f)$ είναι ίσο με $(-1)^n f(x)$ ενώ το ελάχιστο πολυώνυμο του $C(f)$ είναι ίσο με το $f(x)$.

Απόδειξη. Το χαρακτηριστικό πολυώνυμο είναι ίσο με

$$\text{Ch}_{C(f)}(x) = \det \begin{pmatrix} -x & 0 & \cdots & 0 & -a_0 \\ 1 & -x & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & -x & \vdots \\ 0 & \cdots & 0 & 1 & -(x + a_{n-1}) \end{pmatrix}$$

Αν πολλαπλασιάσουμε την τελευταία γραμμή με x και την προσθέσουμε στην προτελευταία στην συνέχεια την προτελευταία στην προηγούμενη της και συνεχίσουμε με τον ίδιο τρόπο μέχρι την δεύτερη γραμμή η ορίζουσα δεν αλλάζει και καταλήγουμε στην ορίζουσα

$$\text{Ch}_{C(f)}(x) = \det \begin{pmatrix} 0 & 0 & \cdots & 0 & f(x) \\ 1 & 0 & \cdots & 0 & * \\ 0 & 1 & \ddots & \vdots & * \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & -(x + a_{n-1}) \end{pmatrix}$$

όπου στην πάνω δεξιά γωνία έχουμε πάρει το

$$f(x) = -(a_0 + x(a_1 + x(\cdots(a_{n-2} + x(a_{n-1} + x))))))$$

μία γραφή του πολυωνύμου x γνωστή και ως σχήμα Horner. Το αποτέλεσμα προκύπτει αναπτύσσοντας κατά Laplace την ορίζουσα κατά μήκος της πρώτης γράμμης.

Για το ελάχιστο πολυώνυμο τώρα: Παρατηρούμε ότι για $x = (x_1, x_2, \dots, x_n)^t$

$$C(f)x = x_n \begin{pmatrix} -a_0 \\ -a_1 \\ \vdots \\ -a_{n-1} \end{pmatrix} + \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{n-2} \end{pmatrix}.$$

Συνεπώς, για $v = (1, 0, \dots, 0)^t$, έχουμε

$$C(f)v = (0, 1, \dots, 0)^t, C(f)^2v = (0, 0, 1, \dots, 0)^t, \dots, C(f)^{n-1}v = (0, \dots, 0, 1)^t.$$

Άρα για μια επιλογή $b_0, b_1, \dots, b_{n-1}$

$$\sum_{j=0}^{n-1} b_j C(f)^j v = (b_0, b_1, \dots, b_{n-1})^t.$$

Το τελευταίο όμως σημαίνει ότι δεν υπάρχει μη-μηδενικό πολυώνυμο $g(x) = \sum_{j=0}^{n-1} b_j x^j$, ώστε $g(C(f)) = 0$, άρα το ελάχιστο πολυώνυμο έχει βαθμό n και ταυτίζεται με το $f(x)$. $\square$

Παρατήρηση VI.2.12. Θα δώσουμε άλλη μια κομψή απόδειξη της πρότασης VI.2.11 η οποία είναι λιγότερο στοιχειώδης από αυτή που ήδη γράψαμε.

Θεωρούμε τον διανυσματικό χώρο $V = \mathbb{F}[x]/\langle f \rangle$. Η συνάρτηση

$$\begin{aligned}\phi : \mathbb{F}[x]/\langle f \rangle &\longrightarrow \mathbb{F}[x]/\langle f \rangle \\ g + \langle f \rangle &\longmapsto x \cdot g + \langle f \rangle\end{aligned}$$

είναι καλά ορισμένη και γραμμική. Ο πίνακας της γραμμικής συνάρτησης της ϕ ως προς την βάση $\{1, x, x^2, \dots, x^{n-1}\}$ είναι ο συνοδός πίνακας του πολυωνύμου $f(x)$. Έστω

$$p_\phi(x) = \sum_{i=1}^d b_i x^i \in \mathbb{F}[x]$$

το ελάχιστο πολυώνυμο της ϕ . Και $\text{Ch}_\phi(x)$ το χαρακτηριστικό. Τότε το p_ϕ είναι η μηδενική συνάρτηση $V \rightarrow V$, δηλαδή

$$\begin{aligned}0 + \langle f \rangle &= p_\phi(\phi)(1 + \langle f \rangle) = \sum_{i=0}^d b_i \phi^i(1 + \langle f \rangle) \\ &= \left(\sum_{i=0}^d a_i x^i \right) + \langle f \rangle = p_\phi + \langle f \rangle\end{aligned}$$

Συνεπώς $f \mid p_\phi \mid \text{Ch}_\phi$. Αφού όλα τα πολυώνυμα είναι μονικά βαθμού n η ισότητα $f(x) = p_\phi(x) = \text{Ch}_\phi(x)$ έπεται.

Παρατήρηση VI.2.13. Παρατηρούμε ότι για μια ρίζα λ του $f(x)$ έχουμε

$$(1, \lambda, \lambda^2, \dots, \lambda^{n-1}) \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix} = \lambda(1, \lambda, \lambda^2, \dots, \lambda^{n-1})$$

Δηλαδή, το $(1, \lambda, \lambda^2, \dots, \lambda^{n-1})^t$ είναι ένα ιδιοδιάνυσμα του $C(f)^t$. Γνωρίζουμε ότι αν το $f(x)$ έχει διαφορετικές ρίζες τότε ο πίνακας $C(f)$ άρα και ο $C(f)^t$ είναι διαγγοποιήσιμοι. Παρατηρούμε ότι σε αυτή την περίπτωση, αν $\lambda_1, \dots, \lambda_n$ είναι οι διαφορετικές ρίζες του $f(x)$ τότε οι στήλες του πίνακα

$$Q = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix}$$

αποτελούν μία βάση από ιδιοδιανύσματα του $C(f)^t$. Παρατηρήστε ότι

$$\det Q = \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j),$$

παράγραφος III.9 συμβατό με το ότι ο Q είναι αντιστρέψιμος.

Εφαρμογή 1. Θα δώσουμε τώρα μια απόδειξη των ταυτοτήτων του Newton (πρόταση [1.8.33](#)) με την βοήθεια της γραμμικής άλγεβρας. Ακολουθούμε τον D. Kalman [\[8\]](#).

Θεωρούμε το πολυώνυμο

$$f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0,$$

με ρίζες $z_1, \dots, z_n$. Σύμφωνα με την πρόταση [1.8.33](#) οι συντελεστές του πολυωνύμου a_i δίνονται από

$$a_i = (-1)^{n-i} e_i(z_1, \dots, z_n)$$

Οπότε αν $p_i(z_1, \dots, z_n) = z_1^i + \dots + z_n^i$, τότε οι ταυτότητες Newton γράφονται ως

$$p_k + a_{n-1}p_{k-1} + \dots + a_0p_{k-n} = 0 \text{ αν } k > n$$

$$p_k + a_{n-1}p_{k-1} + \dots + a_{n-k+1}p_1 = -ka_{n-k} \text{ αν } 1 \leq k \leq n.$$

θεωρούμε τώρα τον $n \times n$ συνοδό πίνακα C του $f(x)$. Αυτός έχει ιδιοτιμές τις ρίζες $z_1, \dots, z_n$ του $f(x)$ ενώ τα αθροίσματα

$$p_k(z_1, \dots, z_n) = \text{tr}(C^k).$$

Συνεπώς, για $k > n$ η πρώτη ταυτότητα Newton γράφεται

$$\text{tr}(C^k) + a_{n-1}\text{tr}(C^{k-1}) + \dots + a_0\text{tr}(C^{k-n}) = 0$$

και επειδή η συνάρτηση ίχνους είναι γραμμική έχουμε ότι η παραπάνω εξίσωση είναι ισοδύναμη με την

$$\text{tr}(C^k + a_{n-1}C^{k-1} + \dots + a_0C^{k-n}) = 0 \Leftrightarrow \text{tr}(C^{k-n}f(C)) = 0$$

η οποία είναι άμεση συνέπεια του θεωρήματος Caley-Hamilton.

Για την περίπτωση $1 \leq k \leq n$ η ταυτότητα Newton γράφεται ως

$$\text{tr}(C^k) + a_{n-1}\text{tr}(C^{k-1}) + \dots + a_{n-k+1}\text{tr}(C) = -ka_{n-k}$$

την οποία την γράφουμε στην ισοδύναμη μορφή:

$$\text{tr}(C^k + a_{n-1}C^{k-1} + \dots + a_{n-k+1}C + a_{n-k}\mathbb{I}_n) = (n-k)a_{n-k}. \quad (\text{VI.7})$$

Γράφουμε $X = x\mathbb{I}_n$ και υπολογίζουμε ότι

$$\begin{aligned} f(X) &= (X - C)(X^{n-1} + (C + a_{n-1}\mathbb{I}_n)X^{n-2} + (C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)X^{n-3} \\ &\quad + \dots + (C^{n-1} + a_{n-1}C^{n-2} + \dots + a_1\mathbb{I}_n)\mathbb{I}_n). \end{aligned}$$

Θα θέλαμε να χρησιμοποιήσουμε την συνάρτηση ίχνους, η οποία όμως δεν συμπεριφέρεται με κάποιο εύκολο τρόπο στα γινόμενα. Παρατηρούμε ότι αν το x δεν είναι ιδιοτιμή του C ο πίνακας $X - C$ είναι αντιστρέψιμος οπότε με αυτή την παραδοχή γράφουμε

$$\begin{aligned} (X - C)^{-1}f(X) &= X^{n-1} + (C + a_{n-1}\mathbb{I}_n)X^{n-2} + (C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)X^{n-3} \\ &\quad + \dots + (C^{n-1} + a_{n-1}C^{n-2} + \dots + a_1\mathbb{I}_n)\mathbb{I}_n. \end{aligned} \quad (\text{VI.8})$$

Στην συνέχεια εφαρμόζουμε τον τύπο του ίχνους για να πάρουμε

$$\begin{aligned} \text{tr}((X - C)^{-1}f(X)) &= nx^{n-1} + \text{tr}(C + a_{n-1}\mathbb{I}_n)x^{n-2} + \text{tr}(C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)x^{n-3} \\ &\quad + \dots + \text{tr}(C^{n-1} + a_{n-1}C^{n-2} + \dots + a_1\mathbb{I}_n). \end{aligned}$$

Θα δείξουμε ότι το αριστερό μέρος της παραπάνω εξίσωσης είναι το $f'(x)$. Με αυτή την παραδοχή, εξισώνοντας τους συντελεστές από το αριστερό μέρος έχουμε το ζητούμενο. Πράγματι, εξισώνοντας τον συντελεστή του x^{n-k-1} στο $p'(x)$ με τον συντελεστή του αριστερού μέρους της εξίσωσης (VI.12) έχουμε

$$(n-k)a_{n-k} = \text{tr}(C^k + a_{n-1}C^{k-1} + \dots + a_{n-k+1}C + a_{n-k}\mathbb{I}_n)$$

η οποία είναι ακριβώς η (VI.7).

Θα υπολογίσουμε τώρα το $A = (X - C)^{-1}f(X)$. Παρατηρούμε ότι $f(X) = f(x\mathbb{I}_n) = f(x)\mathbb{I}_n$, συνεπώς $A = f(x)(x\mathbb{I}_n - C)^{-1}$. Συνεπώς

$$\text{tr}(A) = f(x)\text{tr}(x\mathbb{I}_n - C)^{-1}.$$

Είναι σαφές ότι το ίχνος ενός πίνακα είναι το άθροισμα των ιδιοτιμών του και οι ιδιοτιμές του $(x\mathbb{I}_n - C)^{-1}$ είναι οι $1/(x - z_1), 1/(x - z_2), \dots, 1/(x - z_n)$. Συνεπώς

$$\text{tr}(A) = f(x) \left(\frac{1}{x - z_1} + \frac{1}{x - z_2} + \dots + \frac{1}{x - z_n} \right)$$

το οποίο είναι η παράγωγος του $f(x) = (x - z_1)(x - z_2)\dots(x - z_n)$.

Εφαρμογή 2. Μια απόδειξη του Θεμελιώδους Θεωρήματος της Άλγεβρας

Κάθε απόδειξη θα πρέπει να χρησιμοποιεί μια ιδιότητα των πραγματικών αριθμών. Η ιδιότητα που θα χρησιμοποιήσουμε είναι το γεγονός ότι κάθε πολυώνυμο περιττού βαθμού στο $\mathbb{R}[x]$ έχει μια πραγματική ρίζα και η οποία χρησιμοποιεί το γεγονός ότι ένα τέτοιο πολυώνυμο για αρκετά μεγάλες τιμές του x είναι θετικό για αρκετά μικρές τιμές του x είναι αρνητικό και επειδή η πολυωνυμική συνάρτηση είναι συνεχής θα μηδενιστεί σε κάποιο ενδιάμεσο σημείο. Η απόδειξη που παρουσιάζουμε έχει βασίζεται στην ιδέα του H. Derksen [5] και στην απλοποίηση της από τον K. Conrad [4].

Το θεμελιώδες θεώρημα της Άλγεβρας αναφέρει ότι

Θεώρημα VI.2.14. Κάθε μη σταθερό πολυώνυμο στο $\mathbb{C}[x]$ έχει μια μιγαδική ρίζα.

Θεωρούμε το πολυώνυμο

$$f(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 \in \mathbb{C}[x]$$

το οποίο έχει βαθμό $n \geq 1$. Ο πίνακας

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

σύμφωνα με την πρόταση (VI.2.11) έχει χαρακτηριστικό πολυώνυμο ίσο με το $f(x)$. Συνεπώς το θεώρημα (VI.2.14) είναι ισοδύναμο με το

Θεώρημα VI.2.15. Για κάθε $n \geq 1$ κάθε πίνακας $A \in \mathbb{C}^{n,n}$ έχει ένα ιδιοδιάνυσμα στο $\mathbb{C}^n$. Ισοδύναμα, για κάθε $n \geq 1$, κάθε γραμμική συνάρτηση $L : V \rightarrow V$, με V μιγαδικό διανυσματικό χώρο διάστασης $\dim V = n$, έχει ένα ιδιοδιάνυσμα στο V .

Πράγματι, αν κάθε $n \times n$ πίνακας έχει ένα ιδιοδιάνυσμα τότε και ο συνοδός πίνακας που γράψαμε παραπάνω έχει ένα ιδιοδιάνυσμα δηλαδή το χαρακτηριστικό του πολυώνυμο έχει μια ρίζα. Αντιστρόφως για να δείξουμε ότι ένας πίνακας στο $\mathbb{C}^{n,n}$ έχει ένα ιδιοδιάνυσμα χρησιμοποιήσαμε ότι το χαρακτηριστικό πολυώνυμο έχει μια τουλάχιστον ρίζα.

Βασικό για την απόδειξη είναι το παρακάτω λήμμα, το οποίο θα το χρησιμοποιήσουμε για n δύναμη του 2.

Λήμμα VI.2.16. Έστω $m > 1$ φυσικός αριθμός. Υποθέτουμε ότι για κάθε $\mathbb{F}$ -διανυσματικό χώρο V του οποίου η διάσταση δεν είναι διαιρετή με m , κάθε γραμμική συνάρτηση $L : V \rightarrow V$ έχει ένα ιδιοδιάνυσμα στον V . Τότε κάθε ζευγάρι γραμμικών συναρτήσεων $L_1, L_2 : V \rightarrow V$ έχει ένα κοινό ιδιοδιάνυσμα στο V .

Απόδειξη. Θα κάνουμε επαγωγή στην διάσταση $d = \dim V$, όσο το d διατρέχει τους ακέραιους που δεν είναι διαιρετοί με m . Για $d = 1$ το ζητούμενο είναι τετριμμένο, όλες οι γραμμικές συναρτήσεις αντιμετωπίζονται και κάθε στοιχείο του V είναι ιδιοδιάνυσμα για κάθε γραμμική συνάρτηση $V \rightarrow V$.

Υποθέτουμε ότι $d > 1$ δεν είναι διαιρετό με m και ότι έχουμε αποδείξει το ζητούμενο για όλα τα μικρότερα d τα οποία δεν είναι διαιρετά με m . Θεωρούμε τις γραμμικές συναρτήσεις $L_1, L_2 : V \rightarrow V$ με $L_1 L_2 = L_2 L_1$. Αφού $m \nmid d$ η υπόθεση μας λέει ότι ο L_1 έχει μια ιδιοτιμή λ . Θέτουμε

$$U = \text{Im}(L_1 - \lambda \text{Id}_V), \quad W = \text{Ker}(L_1 - \lambda \text{Id}_V).$$

Και οι δύο παραπάνω χώροι είναι L_1 -αναλλοίωτοι και $\dim W \geq 1$ αφού το λ είναι ιδιοτιμή του L_1 . Οι παραπάνω χώροι είναι και L_2 -αναλλοίωτοι, λόγω της αντιμετάθεσης των L_1, L_2 . Για παράδειγμα αν $u \in U$, έχουμε $u = L_1(v) - \lambda v$ και

$$L_2 u = L_2 L_1(v) - L_2 \lambda v = L_1 L_2(v) - \lambda L_2(v) \in \text{Im}(L_1 - \lambda \text{Id}_V) = U.$$

Επίσης $\dim U + \dim W = d$ δεν είναι διαιρετό με m , συνεπώς ένας τουλάχιστον από τους U, V έχει διάσταση που δεν διαιρείται από το m . Αν ο υπόχωρος U ή V που δεν είναι διαιρετός με m έχει διάσταση μικρότερη του d , τότε έχουμε το ζητούμενο κοινό ιδιοδιάνυσμα από την επαγωγική υπόθεση. Αφού $\dim W \geq 1$ αναγκαστικά $W = V$ και $U = \{0\}$. Τότε όμως κάθε στοιχείο του V είναι ιδιοδιάνυσμα του L_1 και επειδή και ο L_2 έχει κάποιο ιδιοδιάνυσμα, υπάρχει κοινό ιδιοδιάνυσμα των L_1, L_2 . $\square$

Παρατηρούμε ότι το λήμμα δεν λείπει ότι οι L_1 και L_2 έχουν κοινή ιδιοτιμή αλλ'αλλά κοινό ιδιοδιάνυσμα το οποίο μπορεί να αντιστοιχεί σε διαφορετικές ιδιοτιμές των L_1, L_2 .

Παράδειγμα VI.2.17. Για

$$L_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

τους οποίους τους βλέπουμε ως γραμμικές συναρτήσεις $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ έχουμε ότι $L_1 L_2 = L_2 L_1$, ένα κοινό ιδιοδιάνυσμα είναι το $(1, 0, 1)^t$ με ιδιοτιμή 2 για τον L_1 και ιδιοτιμή 3 για τον L_2 .

Η απόδειξη θα δοθεί με επαγωγή στην μεγαλύτερη δύναμη του 2 που διαιρεί το n , δηλαδή γράφουμε $n = 2^k n'$, με $(n', 2) = 1$. Ας ξεκινήσουμε με την περίπτωση $k = 0$, δηλαδή ο n είναι περιττός.

Θεωρούμε ένα $A \in \mathbb{C}^{n,n}$ για να δείξουμε ότι το A έχει ένα ιδιοδιάνυσμα στο $\mathbb{C}^n$ θα σχετίσουμε τον A με κάποιες γραμμικές συναρτήσεις στον

$$H_n = \{M \in \mathbb{C}^{n,n} : M^* = M\}.$$

Παρατηρήστε ότι ο χώρος H_n των ερμητιανών πινάκων είναι ένας $\mathbb{R}$ -διανυσματικός χώρος διάστασης n^2 πάνω από το $\mathbb{R}$. Παρατηρούμε ότι αφού το n είναι περιττός και το n^2 είναι επίσης περιττός. Επίσης παρατηρούμε ότι κάθε $C \in \mathbb{C}^n$ γράφεται ως

$$C = \frac{C + C^*}{2} + i \frac{C - C^*}{2i},$$

όπου τα $(C + C^*)/2$ και $(C - C^*)/2i$ είναι ερμητιανοί πίνακες. Για ένα πίνακα $B \in \mathbb{C}^{n,n}$ παίρνουμε $C = AB$ και έχουμε

$$AB = \frac{AB + B^* A^*}{2} + i \frac{AB - B^* A^*}{2i}.$$

Αν $B \in H_n$ τότε έχουμε

$$AB = \frac{AB + BA^*}{2} + i \frac{AB - BA^*}{2i}.$$

οπότε θεωρούμε τις γραμμικές συναρτήσεις $L_1, L_2 : H_n \rightarrow H_n$ που ορίζονται ως

$$L_1(B) = \frac{AB + BA^*}{2}, \quad L_2(B) = \frac{AB - BA^*}{2i}.$$

Παρατηρούμε ότι $L_1 L_2 = L_2 L_1$ και αφού ο H_n έχει περιττή (πραγματική) διάσταση έχουμε ότι L_1, L_2 έχουν ένα κοινό ιδιοδιάνυσμα στο H_n . Δηλαδή, για ένα μη μηδενικό πίνακα $B \in H_n$ έχουμε

$$L_1(B) = \lambda_1 B, \quad L_2(B) = \lambda_2 B, \quad \text{με } \lambda_1, \lambda_2 \in \mathbb{R}.$$

Συνεπώς

$$AB = L_1(B) + i L_2(B) = (\lambda_1 + i \lambda_2) B.$$

Κάθε μη μηδενική στήλη του B είναι ένα ιδιοδιάνυσμα για τον A . Αυτό ολοκληρώνει την απόδειξη του θεμελιώδους θεωρήματος της άλγεβρας για n περιττό που είναι η βάση της επαγωγής.

Η επαγωγική υπόθεση είναι η εξής: Για κάθε $d \geq 1$ όπου η μέγιστη δύναμη του 2 που διαιρεί το d είναι μικρότερη του 2^k , δηλαδή το 2^k δεν διαιρεί το d , κάθε γραμμική συνάρτηση $L : V \rightarrow V$ έχει ένα ιδιοδιάνυσμα, όπου ο V είναι ένας $\mathbb{C}$ -διανυσματικός χώρος διάστασης d .

Θα θεωρήσουμε τώρα διανυσματικούς χώρους που η διάστασή τους διαιρείται ακριβώς με 2^k . Έστω $n \geq 1$ με $n = 2^k n'$, $(n', 2) = 1$ και $A \in \mathbb{C}^{n,n}$. Θέλουμε να δείξουμε ότι ο A έχει ένα μιγαδικό ιδιοδιάνυσμα. Θεωρούμε τον χώρο

$$\text{Sym}_n(\mathbb{C}) = \{M \in \mathbb{C}^{n,n} : M^t = M\}.$$

Αυτός είναι ένας μιγαδικός διανυσματικός χώρος διάστασης $n(n+1)/2$. Η μεγαλύτερη δύναμη του 2 που διαιρεί την διάσταση αυτή είναι το 2^{k-1} . Αφού το $n(n+1)/2$ δεν διαιρείται με το 2^k κάθε ζευγάρι από γραμμικές συναρτήσεις που αντιμετωπίζονται έχουν ένα κοινό ιδιοδιάνυσμα στο $\text{Sym}_n(\mathbb{C})$. Ορίζουμε $L_1, L_2 : \text{Sym}_n(\mathbb{C}) \rightarrow \text{Sym}_n(\mathbb{C})$ ως

$$L_1(B) = AB + BA^t, \quad L_2(B) = ABA^t.$$

Οι L_1, L_2 αντιμετατίθενται. Συνεπώς, έχουν ένα κοινό ιδιοδιάνυσμα στο $\text{Sym}_n(\mathbb{C})$, κάποιο μη μηδενικό $B \in \text{Sym}_n(\mathbb{C})$ ώστε

$$AB + BA^t = \lambda B, \quad ABA^t = \mu B, \quad \mu \in \mathbb{C}.$$

Πολλαπλασιάζουμε την πρώτη με A και απλοποιούμε με την δεύτερη για να πάρουμε

$$A^2B + \mu B = \lambda AB$$

δηλαδή

$$(A^2 - \lambda A + \mu)B = 0.$$

Κάθε μιγαδικός αριθμός έχει μια μιγαδική τετραγωνική ρίζα, ο τύπος της διακρίνουσας δείχνει ότι κάθε τετραγωνικό πολυώνυμο παραγοντοποιείται σε γραμμικούς παράγοντες. Οπότε γράφουμε

$$z^2 - \lambda z + \mu = (z - \alpha)(z - \beta).$$

Τότε έχουμε ότι $(A - \alpha I_n)(A - \beta I_n) = 0$ και

$$(A - \alpha I_n)(A - \beta I_n)B = 0.$$

Αν $(A - \beta I_n)B = 0$ τότε κάθε μη μηδενική στήλη του B είναι ένα ιδιοδιάνυσμα του A με ιδιοτιμή β . Αν $(A - \beta I_n)B \neq 0$ τότε κάθε μη μηδενική στήλη του $(A - \beta I_n)B$ είναι ένα ιδιοδιάνυσμα του A με ιδιοτιμή α .

Ασκήσεις

Άσκηση VI.14 Δίνονται δύο γραμμικές συναρτήσεις $L, G : V \rightarrow V$ με ελάχιστα πολυώνυμα $p_L(x), p_G(x)$. Αν ο μέγιστος κοινός διαιρέτης $f(x)p_L(x) + g(x)p_G(x) = 1$ δείξτε ότι η $p_L(G) : V \rightarrow V$ είναι ισομορφισμός. Επίσης δείξτε ότι αν $\ker G \neq \{0\}$ τότε η L είναι ισομορφισμός.

VI.3 Κανονική μορφή Jordan

VI.3.1 Ανάλυση σε άθροισμα αναλλοίωτων υπόχωρων

Παρατηρούμε ότι αν ένας διανυσματικός χώρος V είναι ευθύ άθροισμα L -αναλλοίωτων υπόχωρων

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s,$$

δηλαδή $L(W_i) \subset W_i$ για κάθε $1 \leq i \leq s$ τότε μπορούμε να διαλέξουμε βάσεις B_i των χώρων W_i για $1 \leq i \leq s$ ώστε

$$(L, B, B) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_s \end{pmatrix} \quad (\text{VI.9})$$

όπου $A_i = (L|_{W_i}, B_i, B_i)$. Μια τέτοια ανάλυση της γραμμικής συνάρτησης L θα την λέμε ανάλυση κατά μπλόκ της L . Παρατηρούμε ότι μια γραμμική συνάρτηση L (ή ο αντίστοιχος πίνακας) είναι διαγωνοποιήσιμη αν και μόνο αν υπάρχει μια ανάλυση του χώρου V ως ευθύ άθροισμα μονοδιάστατων L -αναλλοίωτων υπόχωρων.

Πρόταση VI.3.1. Το χαρακτηριστικό πολυώνυμο μιας γραμμικής συνάρτησης η οποία δέχεται μια κατά μπλοκ διαγώνια ανάλυση όπως στην εξίσωση (VI.9) είναι το γινόμενο των χαρακτηριστικών πολυωνύμων των περιορισμών δηλαδή

$$\text{Ch}_L(x) = \prod_{i=1}^s \text{Ch}_{L|W_i}(x).$$

Απόδειξη. Το χαρακτηριστικό πολυώνυμο της L είναι το ίδιο με το το χαρακτηριστικό πολυώνυμο του πίνακα της L ως προς οποιαδήποτε βάση B . Υπολογίζουμε ότι

$$\text{Ch}_L(x) = \det \begin{pmatrix} A_1 - x\mathbb{I}_{\dim(W_1)} & 0 & \cdots & 0 \\ 0 & A_2 - x\mathbb{I}_{\dim(W_2)} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_s - x\mathbb{I}_{\dim(W_s)} \end{pmatrix}$$

και το ζητούμενο προκύπτει από την άσκηση III.6. $\square$

Παρατήρηση VI.3.2. Το ελάχιστο πολυώνυμο της L που έχει μια κατά μπλοκ διαγώνια ανάλυση όπως στην εξίσωση (VI.9) δεν είναι κατ'ανάγκη το γινόμενο των ελαχίστων πολυωνύμων των $L|_{W_i}$. Πράγματι $m_{\text{diag}(\lambda, \dots, \lambda)}(x) = x - \lambda$ και όχι $(x - \lambda)^n$.

Λήμμα VI.3.3. Έστω πίνακας

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_s \end{pmatrix}.$$

Το ελάχιστο πολυώνυμο του A είναι το ελάχιστο κοινό πολλαπλάσιο των ελαχίστων πολυωνύμων των πινάκων A_i .

Απόδειξη. Παρατηρούμε ότι για ένα πολυώνυμο $f(x) \in \mathbb{F}[x]$ έχουμε

$$f(A) = \begin{pmatrix} f(A_1) & 0 & \cdots & 0 \\ 0 & f(A_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(A_s) \end{pmatrix}.$$

Άρα το $f(A) = 0$ αν και μόνο αν $f(A_i) = 0$ για κάθε $1 \leq i \leq s$. Συνεπώς, έχουμε ότι $m_{A_i}(x) \mid m_A(x)$ για κάθε $1 \leq i \leq s$ και κάθε κοινό πολλαπλάσιο των $m_{A_i}(x)$ μηδενίζει τον A . Συνεπώς

$$m_A(x) = \text{Ε.Κ.Π}(m_{A_1}(x), \dots, m_{A_s}(x)).$$

$\square$

Λήμμα VI.3.4. Αν $L_i : V \rightarrow V$, $i = 1, 2$ είναι γραμμικές συναρτήσεις και $L_1 L_2 = L_2 L_1$ τότε ο $\text{Ker} L_1$ και ο $\text{Im} L_1$ είναι L_2 -αναλλοίωτοι.

Απόδειξη. Αν $v \in \text{Ker} L_1$ τότε

$$L_1 L_2(v) = L_2 L_1(v) = L_2 0_V = 0_V$$

συνεπώς $L_2(v) \in \text{Ker} L_1$.

Αν $v \in \text{Im} L_1$ τότε υπάρχει $w \in V$, με $L_1 w = v$. Έχουμε ότι

$$L_2 v = L_2 L_1 w = L_1 (L_2 w),$$

άρα και $L_2 v \in \text{Im} L_1$. $\square$

Πρόταση VI.3.5. Αν το ελάχιστο πολυώνυμο $m_L(x)$ της L διασπάται ως γινόμενο $m_L(x) = m_1(x)m_2(x)$ δύο μεταξύ τους πρώτων πολυωνύμων $m_1(x), m_2(x)$ τότε ο V διασπάται σε ένα ευθύ άθροισμα L -αναλλοίωτων υπόχωρων W_1, W_2 , όπου

1. $W_i = \text{Ker} m_i(L) = \{v \in V : m_i(L)v = 0_V\}$
2. Το ελάχιστο πολυώνυμο της $L|_{W_i}$ είναι το $m_i(x)$.

Απόδειξη. Από την εξίσωση (I.10) στην παρατήρηση I.8.7 έχουμε ότι υπάρχουν πολυώνυμα $g_1(x), g_2(x)$ ώστε

$$1 = g_1(x)m_1(x) + g_2(x)m_2(x),$$

συνεπώς

$$\text{Id}_V = g_1(L)m_1(L) + g_2(L)m_2(L) \quad (\text{VI.10})$$

Για κάθε $v \in V$ έχουμε $m_1(L)g_1(L)v \in W_2$ και $m_2(L)g_2(L)v \in W_1$. Επιπλέον, $W_1 \cap W_2 = \{0_V\}$. Πράγματι αν $v \in W_1 \cap W_2$ τότε η σχέση $m_1(L)v = m_2(L)v = 0_V$ μαζί με την εξίσωση (VI.10) επιβάλλουν ότι $v = 0_V$.

Είναι σαφές ότι οι χώροι W_1, W_2 είναι L -αναλλοίωτοι σύμφωνα το λήμμα VI.3.4 αφού η L αντιμετωπίζεται με κάθε πολυώνυμο του L , άρα $Lm_i(L) = m_i(L)L$ για $i = 1, 2$.

Τέλος παρατηρούμε ότι το $m_i(L)$ μηδενίζει εξ ορισμού κάθε διάνυσμα του $W_i = \text{Ker} m_i(L)$, άρα το ελάχιστο πολυώνυμο $m'_i(x)$ του $L|_{W_i}$ διαιρεί το $m_i(x)$. Επίσης, αφού κάθε $v \in V$ γράφεται ως $v = v_1 + v_2$ με $v_i \in W_i$ έχουμε ότι το $m'(x) = m'_1(x)m'_2(x)$ μηδενίζει τον L , αφού

$$m'(L)(v) = m'_1(L)m'_2(L)(v) = m'_1(L)m'_2(L)(v_1) + m'_1(L)m'_2(L)(v_2) = 0_V.$$

Συνεπώς,

$$m_1(x)m_2(x) = m(x) \mid m'_1(x)m'_2(x) \mid m_1(x)m_2(x)$$

άρα $m'_i(x) = m_i(x)$, $i = 1, 2$. □

Επαγωγικά η παραπάνω πρόταση γενικεύεται στην

Πρόταση VI.3.6. Αν το ελάχιστο πολυώνυμο $m_L(x)$ της L διασπάται ως γινόμενο $m_L(x) = m_1(x)m_2(x)\cdots m_s(x)$ s το πλήθος μεταξύ τους πρώτων πολυωνύμων $m_1(x), m_2(x), \dots, m_s(x)$ τότε ο V διασπάται σε ένα ευθύ άθροισμα L -αναλλοίωτων υπόχωρων $W_1, W_2, \dots, W_s$, όπου

1. $W_i = \text{Ker} m_i(L) = \{v \in V : m_i(L)v = 0_V\}$
2. Το ελάχιστο πολυώνυμο της $L|_{W_i}$ είναι το $m_i(x)$.

Θεώρημα VI.3.7. Η γραμμική συνάρτηση $L : V \rightarrow V$, όπου V είναι διανυσματικός χώρος επί του αλγεβρικά κλειστού σώματος $\mathbb{F}$ είναι διαγωνοποίηση αν και μόνο αν όταν το ελάχιστο πολυώνυμο της είναι γινόμενο διαφορετικών μεταξύ τους γραμμικών παραγόντων.

Απόδειξη. Αν η L είναι διαγωνοποίηση και $\dim V = n$, τότε υπάρχει βάση B ώστε $(L, B, B) = \text{diag}(\lambda_1, \dots, \lambda_n)$. Αν $\lambda_{i_1}, \dots, \lambda_{i_s}$ είναι οι διαφορετικές ανά δύο ιδιοτιμές τότε το ελάχιστο πολυώνυμο είναι το

$$m_L(x) = \prod_{v=1}^s (x - \lambda_{i_v}).$$

□

Αντιστρόφως, αν το ελάχιστο πολυώνυμο είναι γινόμενο γραμμικών παραγόντων τότε η πρόταση **VI.3.6** μας δίνει μια διάσπαση του V σε L -αναλλοιώτους υπόχωρους

$$V = W_1 \oplus \cdots \oplus W_s$$

όπου ο περιορισμοί $L|_{W_i}$ έχουν ελάχιστο πολυώνυμο ίσο με $x - \lambda_i$ δηλαδή $L|_{W_i}$ είναι διαγώνιος πίνακας με ίδια ιδιοτιμή λ_i .

Λήμμα VI.3.8. Αν $L : V \rightarrow V$ είναι διαγωνοποιήσιμη γραμμική συνάρτηση και $W \subset V$ είναι L -αναλλοιώτος υπόχωρος τότε $L|_W$ είναι διαγωνοποιήσιμη.

Απόδειξη. Το ελάχιστο πολυώνυμο $m_{L|_W}(x)$ της $L|_W$ διαιρεί το ελάχιστο πολυώνυμο $m_L(x)$ της L . Πράγματι, αφού $m_L(L)(v) = 0_V$ για κάθε $v \in V$, το ίδιο θα ισχύει και για κάθε $w \in W$, δηλαδή $m_L(L|_W) = 0_W$. Όμως η L είναι διαγωνοποιήσιμη και συνεπώς το ελάχιστο πολυώνυμο της διασπάται σε γραμμικούς, διαφορετικούς ανά δύο γραμμικούς παράγοντας. Το ίδιο θα συμβαίνει και με το $m_{L|_W}(x)$ και συνεπώς και η $m_{L|_W}$ είναι διαγωνοποιήσιμη. $\square$

Πρόταση VI.3.9. Θεωρούμε δύο διαγωνοποιήσιμες γραμμικές συναρτήσεις $L_1, L_2 : V \rightarrow V$. Υπάρχει κοινή βάση B ώστε $(L_1, B, B), (L_2, B, B)$ διαγώνιοι πίνακες αν και μόνο αν $L_1 L_2 = L_2 L_1$.
Ισοδύναμα, για δύο διαγωνοποιήσιμους πίνακες $A_1, A_2 \in \mathbb{F}^{n,n}$ υπάρχει κοινός αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{n,n}$ αν και μόνο αν $A_1 A_2 = A_2 A_1$.

Απόδειξη. Αν υπάρχει τέτοια βάση B , ώστε $(L_1, B, B), (L_2, B, B)$ διαγώνιοι, τότε $(L_1, B, B)(L_2, B, B) = (L_1, B, B)L_2, B, B)$ γιατί οι διαγώνιοι πίνακες αντιμετατίθενται. Συνεπώς $L_1 L_2 = L_2 L_1$.

Αντιστρόφως, έστω $L_1 L_2 = L_2 L_1$. Αφού η L_1 είναι διαγωνοποιήσιμη, έχει μια βάση αποτελούμενη από ιδιοδιανύσματα, δηλαδή

$$V = \text{Ker}(L_1 - \lambda_1 I_V) \oplus \cdots \oplus \text{Ker}(L_1 - \lambda_s I_V)$$

όπου $\lambda_1, \dots, \lambda_s$ είναι οι ανά δύο διαφορετικές ιδιοτιμές της L_1 . Σύμφωνα με το λήμμα **VI.3.4** οι ιδιόχωροι $\text{Ker}(L_1 - \lambda_i \text{Id}_V)$ είναι L_2 -αναλλοιώτοι, και από το λήμμα **VI.3.8** οι $L_2|_{\text{Ker}(L_1 - \lambda_i \text{Id}_V)}$ είναι διαγωνοποιήσιμη. Αυτό σημαίνει ότι ο ιδιόχωρος $\text{Ker}(L_1 - \lambda_i \text{Id}_V)$ της ιδιοτιμής λ_i , ο οποίος αποτελείται από ιδιοδιανύσματα της L_1 ως προς την ιδιοτιμή λ_i έχει μια βάση από ιδιοδιανύσματα της L_2 . Δηλαδή, υπάρχει μια βάση που αποτελείται ταυτόχρονα από ιδιοδιανύσματα και για την L_1, L_2 , δηλαδή το ζητούμενο.

Η ταυτόχρονη διαγωνοποίηση των πινάκων είναι ισοδύναμη με την ταυτόχρονη διαγωνοποίηση των αντίστοιχων γραμμικών συναρτήσεων. $\square$

Ορισμός VI.3.10. Μια γραμμική συνάρτηση $L : V \rightarrow V$ θα λέγεται ημιαπλή αν και μόνο αν για κάθε L -αναλλοίωτο υπόχωρο $W \subset V$ υπάρχει L -αναλλοίωτος υπόχωρος W' ώστε $V = W \oplus W'$.

Πρόταση VI.3.11. Μια γραμμική συνάρτηση $L : V \rightarrow V$, όπου V είναι διανυσματικός χώρος επί του αλγεβρικά κλειστού σώματος $\mathbb{F}$, είναι ημιαπλή αν και μόνο αν είναι διαγωνοποιήσιμη.

Απόδειξη. Αν η L είναι διαγωνοποιήσιμη τότε σύμφωνα με το λήμμα VI.3.8 και ο περιορισμός της στο W είναι διαγωνοποιήσιμος. Διαλέγουμε μια βάση του W την $w_1, \dots, w_s$ αποτελούμενη από ιδιοδιανύσματα την οποία την συμπληρώνουμε σε μία βάση από ιδιοδιανύσματα $w'_1, \dots, w'_r$ του V . Αυτό μπορεί να γίνει με την βοήθεια της πρότασης IV.3.22. Το ζητούμενο αναλλοίωτο συμπλήρωμα είναι το $W' = \langle w'_1, \dots, w'_r \rangle$.

Για το αντίστροφο θα χρησιμοποιήσουμε επαγωγή. Έστω ότι κάθε ημιαπλή γραμμική συνάρτηση $L' : W \rightarrow W$ είναι διαγωνοποιήσιμη για κάθε διανυσματικό χώρο διάστασης $n-1$. Θεωρούμε την $L : V \rightarrow V$, όπου $\dim V = n$. Αφού το $\mathbb{F}$ είναι αλγεβρικά κλειστό η L έχει μια ιδιοτιμή και έναν L -αναλλοίωτο μονοδιάστατο χώρο $\langle v \rangle$ που αντιστοιχεί σε ένα ιδιοδιάνυσμα $v \neq 0$ της ιδιοτιμής. Η ημιαπλότητα της L μας δίνει την ύπαρξη ενός L -αναλλοίωτου W' , ώστε $V = \langle v \rangle \oplus W'$. Το ζητούμενο έπεται από την επαγωγική υπόθεση αφού $\dim W' = n-1$. $\square$

Παρατήρηση VI.3.12. Έστω $L : V \rightarrow V$ και έστω W ένας L -αναλλοίωτος υπόχωρος. Μπορούμε να ορίζουμε μια γραμμική συνάρτηση $L_{V/W} : V/W \rightarrow V/W$, ως εξής:

$$V/W \ni v/w \mapsto L_{V/W}(v + W) = L(v) + W \in V/W.$$

Η παραπάνω είναι καλά ορισμένη αφού αν $v = v' + w$, έχουμε ότι $L(v) = L(v') + L(w)$, με $L(w) \in W$. Αν διαλέξουμε μία βάση $B_W = \{w_1, \dots, w_n\}$ του W την οποία την επεκτείνουμε σε βάση $B_V = \{w_1, \dots, w_n, v_1, \dots, v_s\}$ του V τότε ο πίνακας

$$(L, B_V, B_V) = \left(\begin{array}{c|c} (L|_W, B_W, B_W) & * \\ \hline \mathbf{0}_{s,n} & (L_{V/W}, B_{V/W}, B_{V/W}) \end{array} \right), \quad (\text{VI.11})$$

όπου $B_{V/W} = \{v_1 + W, \dots, v_s + W\}$ είναι μια βάση του V/W .

Πρόταση VI.3.13. Τα χαρακτηριστικά πολυώνυμα $\text{Ch}_L(x)$, $\text{Ch}_{L|_W}(x)$ και $\text{Ch}_{L_{V/W}}(x)$ συνδέονται με την σχέση

$$\text{Ch}_L(x) = \text{Ch}_{L|_W}(x) \cdot \text{Ch}_{L_{V/W}}(x).$$

Απόδειξη. Από την εξίσωση (VI.11) έχουμε ότι

$$\text{Ch}_L(x) = \det \left(\begin{array}{c|c} (L|_W, B_W, B_W) - x\mathbb{I}_n & * \\ \hline \mathbf{0}_{s,n} & (L_{V/W}, B_{V/W}, B_{V/W}) - x\mathbb{I}_s \end{array} \right) = \text{Ch}_{L|_W}(x) \cdot \text{Ch}_{L_{V/W}}(x).$$

$\square$

VI.3.2 Τριγωνοποιήσιμοι Πίνακες

Θα αποδείξουμε μια

Θεώρημα VI.3.14. Θεωρούμε μια γραμμική συνάρτηση $L : V \rightarrow V$, όπου ο V έχει πεπερασμένη διάσταση n , πάνω από ένα αλγεβρικά κλειστό σώμα $\mathbb{F}$. Υπάρχει μια βάση $B = \{v_1, \dots, v_n\}$ του V στην οποία ο πίνακας (L, B, B) είναι άνω τριγωνικός δηλαδή της μορφής

$$(L, B, B) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Απόδειξη. Αν $\dim V = 1$ το αποτέλεσμα είναι σαφές. Θα αποδείξουμε το ζητούμενο με επαγωγή. Υποθέτουμε ότι το θεώρημα ισχύει για κάθε διανυσματικό χώρο διάστασης $n - 1$.

Το χαρακτηριστικό πολυώνυμο του έχει βαθμό n και συνεπώς έχει μια ρίζα που είναι ιδιοτιμή και έχει ένα μη-μηδενικό ιδιοδιάνυσμα v_1 . Συμπληρώνουμε σε μία βάση $B_1 = \{v_1, b_2, \dots, b_n\}$ του V ως προς την οποία ο πίνακας της L γράφεται

$$(L, B_1, B_1) = \left(\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{n-1,n} \\ 0 & a_{n2} & \cdots & a_{nn} \end{array} \right) = \left(\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} (L_{V/\langle v_1 \rangle}, B'_1, B'_1) \end{array} \right),$$

όπου $L_{V/\langle v_1 \rangle}$ είναι η επαγόμενη συνάρτηση στον χώρο πηλίκο $V/\langle v_1 \rangle$ και $B_1 = \{b_2 + \langle v_1 \rangle, \dots, b_n + \langle v_1 \rangle\}$ είναι μια βάση του χώρου πηλίκου. Παρατηρούμε ότι $\dim V/\langle v_1 \rangle = n - 1$ οπότε από την επαγωγική υπόθεση έχουμε ότι υπάρχει μια βάση $B'_1 = \{v_2 + \langle v_1 \rangle, \dots, v_n + \langle v_1 \rangle\}$ ως προς την οποία ο $(L_{V/\langle v_1 \rangle}, B'_1, B'_1)$ να είναι άνω τριγωνικός. Ο πίνακας της L ως προς την βάση $B = \{v_1, v_2, \dots, v_n\}$, έχει την ζητούμενη μορφή. $\square$

Παρατηρούμε ότι τα στοιχεία της διαγωνίου είναι ιδιοτιμές, δηλαδή $a_{ii} = \lambda_i$. Έχουμε το χαρακτηριστικό πολυώνυμο

$$\text{Ch}_A(x) = (-1)^n \prod_{i=1}^n (x - \lambda_i) = (-1)^n \left(x^n - \sum_{i=1}^n x^{n-1} \lambda_i + \cdots + (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n \right).$$

Δηλαδή ο σταθερός όρος του χαρακτηριστικού πολυωνύμου είναι το γινόμενο των ιδιοτιμών δηλαδή η ορίζουσα, κάτι το οποίο γνωρίζαμε και από τον ίδιο τον ορισμό του χαρακτηριστικού πολυωνύμου. Αλλά το άθροισμα των ιδιοτιμών είναι το ίχνος του πίνακα και σχετίζεται με τον συντελεστή a_{n-1} του x^{n-1} στο χαρακτηριστικό πολυώνυμο από την σχέση

$$\text{tr}(A) = (-1)^{n+1} a_{n-1}.$$

Παρατηρήστε ότι όμοιοι πίνακες έχουν το ίδιο ίχνος, άρα το ίχνος μπορεί να υπολογιστεί από το άθροισμα των ιδιοτιμών, δείτε και την άσκηση [II.6](#).

Ασκήσεις

Άσκηση VI.15 Δίνεται ένας 3×3 πίνακας με συντελεστές πραγματικούς. Δείξτε ότι αν ο πίνακας δεν είναι τριγωνοποιήσιμος στο $\mathbb{R}$ τότε είναι διαγωνοποιήσιμος στο $\mathbb{C}$.

Άσκηση VI.16 Θεωρούμε τον χώρο των πολυωνύμων $\mathbb{R}_2[x]$ βαθμού το πολύ ίσου με δύο και την γραμμική απεικόνιση:

$$f : \mathbb{R}_2[x] \rightarrow \mathbb{R}_2[x]$$

που ορίζεται ως:

$$\begin{aligned} f(x^2 + x + 1) &= x^2 + x + 1 \\ f(x) &= ax^2 + bx \\ f(1) &= 0 \end{aligned}$$

με $a \neq b$.

1. Δείξτε ότι η f είναι τριγωνίσιμη για όλες τις τιμές $a, b \in \mathbb{R}$.
2. Εξετάστε για ποιες τιμές των a, b , είναι η f διαγωνίσιμη.

VI.3.3 Ανάλυση Jordan

Ορισμός VI.3.15. Θα ονομάζουμε *Jordan block* $J(\lambda, k)$ τον $k \times k$ πίνακα

$$J(\lambda, k) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

Λήμμα VI.3.16. Το χαρακτηριστικό και ελάχιστο πολυώνυμο του $J(\lambda, k)$ υπολογίζονται

$$\begin{aligned} \text{Ch}_{J(\lambda, k)}(x) &= (-1)^k (x - \lambda)^k \\ p_{J(\lambda, k)}(x) &= (x - \lambda)^k \end{aligned}$$

Απόδειξη. Το χαρακτηριστικό πολυώνυμο προκύπτει από τον υπολογισμό της ορίζουσας

$$\text{Ch}_{J(\lambda, k)}(x) = \det(J(\lambda, k) - x\mathbb{I}_k) = \det \begin{pmatrix} \lambda - x & 1 & 0 & \cdots & 0 \\ 0 & \lambda - x & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda - x & 1 \\ 0 & \cdots & \cdots & 0 & \lambda - x \end{pmatrix} = (\lambda - x)^k.$$

Το ελάχιστο πολυώνυμο $p_{J(\lambda, k)}$ διαιρεί το χαρακτηριστικό πολυώνυμο συνεπώς είναι της μορφής $(x - \lambda)^\nu$ για κάποιο $0 < \nu \leq k$. Από το παράδειγμα VI.2.5 έχουμε ότι $\nu = k$. $\square$

Παρατήρηση VI.3.17. Ο πίνακας $J(\lambda, k)$ για $k \geq 2$ δεν είναι διαγωνοποιήσιμος. Πράγματι η αλγεβρική πολλαπλότητα της ιδιοτιμής λ είναι ίση με k ενώ η γεωμετρική ίση με 1.

Θεώρημα VI.3.18 (Κανονική μορφή Jordan). Για ένα $\mathbb{F}$ -διανυσματικό χώρο όπου το $\mathbb{F}$ είναι αλγεβρικά κλειστό σώμα, οποιοσδήποτε πίνακας $A \in \mathbb{F}^{n, n}$ είναι όμοιος με τον πίνακα

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_s \end{pmatrix} \quad (\text{VI.12})$$

όπου για κάθε $1 \leq i \leq s$ οι πίνακες $J_i = J(\lambda_i, k_i)$ για $\lambda_i \in \mathbb{F}$ και $0 < k_i \in \mathbb{N}$.

Παρατήρηση VI.3.19. Στην παραπάνω ανάλυση οι τιμές $\lambda_1, \dots, \lambda_s$ δεν είναι ανάγκη να είναι διαφορετικές.

Η απόδειξη του θεωρήματος της κανονικής μορφής Jordan βασίζεται στα παρακάτω λήμματα:

Λήμμα VI.3.20. Έστω V ένας $\mathbb{F}$ -διανυσματικός χώρος, όπου το σώμα $\mathbb{F}$ είναι αλγεβρικά κλειστό και $L : V \rightarrow V$ γραμμική συνάρτηση. Έστω $\lambda_1, \dots, \lambda_r$ διαφορετικές ανά δύο ιδιοτιμές της L . Τότε υπάρχουν ακέραιοι $s_1, \dots, s_r$, ώστε

$$V = \text{Ker}(L - \lambda_1 \text{Id}_V)^{s_1} \oplus \text{Ker}(L - \lambda_2 \text{Id}_V)^{s_2} \oplus \cdots \oplus \text{Ker}(L - \lambda_r \text{Id}_V)^{s_r},$$

όπου με $(L - \lambda_v \text{Id}_V)^{s_v}$ στον παραπάνω τύπο εννοούμε την s_v φορές διαδοχική σύνθεση της γραμμικής συνάρτησης $L - \lambda_v \text{Id}_V$.

Απόδειξη. Διαλέγουμε μία από τις ιδιοτιμές του L , έστω λ . Η απόδειξη του λήμματος θα γίνει σε 5 βήματα.

Βήμα 1. Ορίζουμε τον χώρο $W_i = \text{Ker}(L - \lambda \text{Id}_V)^i$ για κάθε $i \geq 1$ οι όποιοι έχουν την παρακάτω σχέση εγκλεισμού

$$W_1 \subset W_2 \subset W_3 \subset \dots \subset W_i \subset \dots$$

Αφού ο χώρος V είναι πεπερασμένης διάστασης η παραπάνω ακολουθία δεν μπορεί να είναι γνήσια αύξουσα για κάθε i , άρα υπάρχει ένα t ώστε $W_t = W_{t+1}$. Υποθέτουμε ότι αυτό είναι το μικρότερο δυνατό t . Είναι σαφές, ότι αν $W_t = W_{t+1}$ τότε $W_t = W_{t+1} = W_{t+2} = \dots$ δηλαδή η ακολουθία των χώρων γίνεται σταθερή. Πράγματι $W_{t+1} \subset W_{t+2}$ και αν $v \in W_{t+2} = \text{Ker}(L - \lambda \text{Id}_V)^{t+2}$ τότε $(L - \lambda \text{Id}_V)v \in \text{Ker}(L - \lambda \text{Id}_V)^{t+1} = \text{Ker}(L - \lambda \text{Id}_V)^t$, άρα $(L - \lambda \text{Id}_V)^t(L - \lambda \text{Id}_V)v = \mathbf{0}_V$ οπότε $v \in \text{Ker}(L - \lambda \text{Id}_V)^{t+1}$.

Βήμα 2. Τώρα θα αποδείξουμε ότι $\text{Ker}(L - \lambda \text{Id}_V)^t \cap \text{Im}(L - \lambda \text{Id}_V)^t = \mathbf{0}_V$. Έστω ότι για ένα $v \in V$ ανήκει στην τομή. Συνεπώς υπάρχει ένα $w \in V$ με $v = (L - \lambda \text{Id}_V)^t w$. Από την άλλη $(L - \lambda \text{Id}_V)^t v = \mathbf{0}_V$, οπότε $(L - \lambda \text{Id}_V)^{2t} w = (L - \lambda \text{Id}_V)^t v = \mathbf{0}_V$, άρα $w \in \text{Ker}(L - \lambda \text{Id}_V)^{2t} = W_{2t} = W_t$, το οποίο δίνει ότι $v = (L - \lambda \text{Id}_V)^t w = \mathbf{0}_V$.

Βήμα 3. Γνωρίζουμε ότι $\dim V = \dim \text{Ker}(L - \lambda \text{Id}_V)^t + \dim \text{Im}(L - \lambda \text{Id}_V)^t$. Επιπλέον γνωρίζουμε ότι

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Οπότε για $V_1 = \text{Ker}(L - \lambda \text{Id}_V)^t$, $V_2 = \text{Im}(L - \lambda \text{Id}_V)^t$ που έχουν μηδενική τομή έχουμε

$$\dim(\text{Ker}(L - \lambda \text{Id}_V)^t + \text{Im}(L - \lambda \text{Id}_V)^t) = \dim V$$

άρα

$$V = \text{Ker}(L - \lambda \text{Id}_V)^t \oplus \text{Im}(L - \lambda \text{Id}_V)^t.$$

Βήμα 4. Παρατηρούμε ότι οι χώροι $\text{Ker}(L - \lambda \text{Id}_V)^t$ και $\text{Im}(L - \lambda \text{Id}_V)^t$ είναι και οι δύο L -αναλλοίωτοι. Πράγματι, παρατηρούμε ότι

$$L(L - \lambda \text{Id}_V)^t = (L - \lambda \text{Id}_V)^t L$$

άρα αν $v \in \text{Ker}(L - \lambda \text{Id}_V)^t$ τότε

$$(L - \lambda \text{Id}_V)^t L(v) = L(L - \lambda \text{Id}_V)^t v = \mathbf{0}_V,$$

οπότε $L \text{Ker}(L - \lambda \text{Id}_V)^t \subset \text{Ker}(L - \lambda \text{Id}_V)^t$. Ομοίως αν $v \in \text{Im}(L - \lambda \text{Id}_V)^t$, τότε $v = (L - \lambda \text{Id}_V)^t w$, οπότε

$$Lv = L(L - \lambda \text{Id}_V)^t w = (L - \lambda \text{Id}_V)^t Lw,$$

άρα $L(\text{Im}(L - \lambda \text{Id}_V)^t) \subset \text{Im}(L - \lambda \text{Id}_V)^t$.

Βήμα 5. Το λήμμα μπορεί να αποδειχθεί με επαγωγή στο πλήθος των διαφορετικών ιδιοτιμών της συνάρτησης L . Αν λοιπόν $\lambda_1, \dots, \lambda_s$ είναι οι διαφορετικές ανά δύο ιδιοτιμές της L , εφαρμόζουμε τα βήματα 1-4 για κάθε $\lambda = \lambda_1$ και στην συνέχεια έχουμε τον περιορισμό της L στον $\text{Im}(L - \lambda_1 \text{Id}_V)^t$ ο οποίος έχει ιδιοτιμές $\lambda_2, \dots, \lambda_s$, οπότε από την επαγωγική υπόθεση έχουμε ότι

$$\text{Im}(L - \lambda_1 \text{Id}_V)^t = \text{Ker}(L - \lambda_2 \text{Id}_V)^{s_2} \oplus \dots \oplus \text{Ker}(L - \lambda_s \text{Id}_V)^{s_r}.$$

□

Λήμμα VI.3.21 (Mark Wildon). Έστω V είναι ένας $\mathbb{F}$ -διανυσματικός χώρος διάστασης n , και $T : V \rightarrow V$ είναι γραμμική συνάρτηση με $T^s = 0_{\text{Env}(V)}$ για κάποιο φυσικό αριθμό s , τότε υπάρχουν διανύσματα $u_1, \dots, u_k$ και φυσικοί αριθμοί $a_1, \dots, a_k$ ώστε

$$T^{a_i}(u_i) = 0_V, \text{ για } i = 1, \dots, k,$$

και τα διανύσματα

$$u_1, T(u_1), \dots, T^{a_1-1}u_1, \dots, u_k, T(u_k), \dots, T^{a_k-1}(u_k)$$

σχηματίζουν μια βάση του χώρου V .

Απόδειξη. Αν η T είναι η μηδενική συνάρτηση, τότε μπορούμε να διαλέξουμε τα $u_1, \dots, u_k$ να είναι μια βάση του V και $a_1 = \dots = a_k = 1$.

Θα δώσουμε μια επαγωγική απόδειξη στην διάσταση του V . Αν το $n = 1$ τότε T^s είναι η μηδενική συνάρτηση αν και μόνο αν η T είναι η μηδενική συνάρτηση οπότε επαναλαμβάνουμε το επιχείρημα για μηδενική συνάρτηση.

Υποθέτουμε ότι το λήμμα ισχύει για διαστάσεις μικρότερες του n και θα το αποδείξουμε για την διάσταση n . Θεωρούμε τον χώρο $\text{Im}T$. Αν $\dim \text{Im}T = 0$ τότε η T είναι η μηδενική συνάρτηση και το λήμμα έπεται. Αν $\dim \text{Im}T = n$ τότε η T είναι 1-1 και δεν μπορεί T^s να είναι η μηδενική συνάρτηση για κάποιο s , άτοπο. Άρα μπορούμε να υποθέσουμε ότι $0 < \dim \text{Im}T < n$, και από την επαγωγική υπόθεση έχουμε ότι υπάρχουν διανύσματα $v_1, \dots, v_\ell$ και φυσικοί αριθμοί $b_1, \dots, b_\ell$ ώστε

$$T^{b_i}(v_i) = 0_V \text{ για } i = 1, \dots, \ell$$

και

$$v_1, T(v_1), \dots, T^{b_1-1}(v_1), \dots, v_\ell, T(v_\ell), \dots, T^{b_\ell-1}(v_\ell) \quad (\text{VI.13})$$

να αποτελούν μια βάση του $\text{Im}T$.

Αφού τα $v_1, \dots, v_\ell \in \text{Im}(T)$ υπάρχουν διανύσματα $w_1, \dots, w_\ell$ ώστε $T(w_i) = v_i$ για $i = 1, \dots, \ell$. Από την άλλη τα διανύσματα $T^{b_1-1}v_1, \dots, T^{b_\ell-1}v_\ell$ είναι γραμμικά ανεξάρτητα διανύσματα του $\text{Ker}T$, τα οποία τα επεκτείνουμε σε μία βάση

$$T^{b_1-1}v_1, \dots, T^{b_\ell-1}v_\ell, z_1, \dots, z_m \quad (\text{VI.14})$$

του $\text{Ker}T$. Παρατηρούμε ότι $T^j(w_i) = T^{j-1}(v_i)$ για όλες τις επιλογές των i, j . Αρκεί να αποδείξουμε ότι

$$w_1, T(w_1), \dots, T^{b_1}(w_1), \dots, w_\ell, T(w_\ell), \dots, T^{b_\ell}(w_\ell), z_1, \dots, z_m \quad (\text{VI.15})$$

αποτελούν μια βάση του V .

Θα δείξουμε πρώτα ότι είναι γραμμικά ανεξάρτητα. Υποθέτουμε την ύπαρξη ενός γραμμικού συνδυασμού

$$\sum_{i=1}^{\ell} \sum_{j=0}^{b_i} \alpha_{i,j} T^j(w_i) + \sum_{v=1}^m \beta_v z_v = 0_V. \quad (\text{VI.16})$$

Εφαρμόζουμε τον T στην παραπάνω σχέση για να πάρουμε

$$\sum_{i=1}^{\ell} \sum_{j=0}^{b_i} \alpha_{i,j} T^{j+1}(w_i) = \sum_{i=1}^{\ell} \sum_{j=0}^{b_i-1} \alpha_{i,j} T^j(v_i)$$

συνεπώς τα

$$\alpha_{1,0} = \alpha_{1,1} = \alpha_{1,b_1-1} = \dots = \alpha_{\ell,0} = \alpha_{\ell,1} = \dots = \alpha_{\ell,b_\ell-1} = 0.$$

Επιστρέφουμε με αυτή την πληροφορία στην εξίσωση (VI.16) για να πάρουμε

$$\alpha_{1,b_1} T^{b_1}(w_1) + \dots + \alpha_{\ell,b_\ell} T^{b_\ell}(w_\ell) + \beta_1 z_1 + \dots + \beta_m z_m = 0_V,$$

όμως τα διανύσματα που εμφανίζονται στο παραπάνω άθροισμα είναι βάση του $\text{Ker} T$, οπότε

$$\alpha_{1,b_1} = \dots = \alpha_{\ell,b_\ell} = \beta_1 = \dots = \beta_m = 0.$$

Για να δείξουμε ότι το σύνολο των διανυσμάτων της εξίσωσης (VI.15) είναι βάση αρκεί, σύμφωνα με το θεώρημα IV.3.11, να δείξουμε ότι το πλήθος τους είναι όση και η διάσταση του χώρου V . Τα στοιχεία της εξίσωσης (VI.13) είναι μια βάση του $\text{Im} T$, οπότε $\dim \text{Im} T = b_1 + \dots + b_\ell$. Τα διανύσματα της εξίσωσης (VI.14) είναι μια βάση του πυρήνα οπότε $\dim \text{Ker} T = l + m$. Συνεπώς

$$\begin{aligned} \dim V &= \dim \text{Ker} T + \dim \text{Im} T = b_1 + \dots + b_\ell + l + m \\ &= (1 + b_1) + \dots + (1 + b_\ell) + m, \end{aligned}$$

δηλαδή ακριβώς το πλήθος των διανυσμάτων της εξίσωσης (VI.15). $\square$

Παράδειγμα VI.3.22. Θεωρούμε τον πίνακα

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Έχουμε ότι $\text{Im}(A) = \langle (1, 0, 0)^t \rangle$. Διαλέγουμε $v_1 = (1, 0, 0)^t$. Τα διανύσματα $(0, 1, 0)^t$ και $(0, 0, 1)^t$ απεικονίζονται μέσω της A στο v_1 , όπως και κάθε διάνυσμα της μορφής $(x_1, x_2, 1 - x_2)^t$. Διαλέγουμε ένα από αυτά, για παράδειγμα το $u_1 = (0, 0, 1)^t$. Τέλος, συμπληρώνουμε την βάση με ένα διάνυσμα του πυρήνα για παράδειγμα το $u_2 = (0, 1, -1)^t$. Έτσι έχουμε την βάση $\{u_1, T(u_1) = v_1, u_2\}$, οπότε $a_1 = 2, a_2 = 1$ και παρατηρούμε ότι $T^2(u_1) = 0 = T(u_2)$.

Προχωράμε τώρα στην απόδειξη του θεωρήματος VI.3.18. Από το λήμμα VI.3.20 έχουμε ότι υπάρχουν ακέραιοι $s_1, \dots, s_r$ ώστε

$$V = \text{Ker}(L - \lambda_1 \text{Id}_V)^{s_1} \oplus \dots \oplus \text{Ker}(L - \lambda_r \text{Id}_V)^{s_r},$$

όπου $\lambda_1, \dots, \lambda_r$ είναι οι διαφορετικές ιδιοτιμές του L . Από το λήμμα VI.3.21 διαλέγουμε βάσεις για τους υπόχωρους $\text{Ker}(L - \lambda_1 \text{Id}_V), \dots, \text{Ker}(L - \lambda_r \text{Id}_V)$ της οποίες τις γράφουμε την μία μετά την άλλη και παρατηρούμε ότι ο πίνακας της γραμμικής συνάρτησης ως προς αυτή την βάση είναι πράγματι ο πίνακας της εξίσωσης (VI.12).

Μέθοδος 4 (Υπολογισμός κανονικής μορφής Jordan). Θεωρούμε μια γραμμική συνάρτηση $L : V \rightarrow V$, η οποία εκφράζεται με την βοήθεια πίνακα A . Για να υπολογίσουμε την κανονική μορφή Jordan, εκτελούμε τα παρακάτω βήματα:

- 1 Υπολογίζουμε τις ιδιοτιμές του πίνακα A και τα ιδιοδιανύσματα που αντιστοιχούν σε κάθε ιδιοτιμή.
- 2 Το πλήθος των γραμμικά ανεξάρτητων ιδιοδιανυσμάτων που υπολογίσαμε δίνει το πλήθος των διαφορετικών Jordan blocks.
- 3 Για κάθε ιδιοδιάνυσμα v που υπολογίσαμε, και που αντιστοιχεί στην ιδιοτιμή λ , λύνουμε το σύστημα $(A - \lambda \mathbb{I}_n)w_1 = v$. Στην συνέχεια υπολογίζουμε το $(A - \lambda \mathbb{I}_n)w_2 = w_1$ και συνεχίζουμε με τον ίδιο τρόπο υπολογίζοντας

$$(A - \lambda \mathbb{I}_n)w_3 = w_2, (A - \lambda \mathbb{I}_n)w_4 = w_3, \dots, (A - \lambda \mathbb{I}_n)w_n = w_{n-1},$$

και η διαδικασία σταματάει όταν το σύστημα $(A - \lambda \mathbb{I}_n)w_{n+1} = w_n$ είναι αδύνατο.

4 Σχηματίζουμε τον πίνακα P ο οποίος έχει ως στήλες τα διανύσματα $v, w_1, \dots, w_n$ τα οποία υπολογίσαμε για κάθε ιδιοτιμή. Ο πίνακας $P^{-1}AP$ είναι ο πίνακας της κανονικής μορφής Jordan.

Παραδείγματα VI.3.23. 1. Θεωρούμε τον πίνακα

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

ο οποίος έχει χαρακτηριστικό πολυώνυμο $(x-2)^3$ και ιδιοτιμή $\lambda = 2$. Το σύστημα $(A - 2I_n)x = 0$ έχει μονοδιάστατο πυρήνα ο οποίος παράγεται από το $v = (1, 0, 0)^t$. Στην συνέχεια λύνουμε το σύστημα

$$(A - 2I_n)w_1 = v \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

το οποίο έχει χώρο λύσεων $x(1, 0, 0)^t + (0, 1, 0)^t, t \in \mathbb{R}$. Διαλέγουμε την απλούστερη λύση $w_1 = (0, 1, 0)^t$. Επαναλαμβάνουμε την διαδικασία

$$(A - 2I_n)w_2 = w_1 \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

το οποίο έχει χώρο λύσεων $x(1, 0, 0)^t + (0, 0, 1)^t, t \in \mathbb{R}$. Διαλέγουμε ως λύση το $w_2 = (0, 0, 1)^t$. Τα διανύσματα v, w_1, w_2 είναι γραμμικά ανεξάρτητα και αποτελούν μια βάση του $\mathbb{R}^3$ οπότε μπορούμε να σταματήσουμε εδώ. Παρατηρούμε ότι στην επόμενη επανάληψη

$$(A - 2I_n)w_3 = w_2 \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

το σύστημα είναι αδύνατο οπότε δεν υπάρχει άλλο w_3 . Ο πίνακας P που αποτελείται από τα διανύσματα v, w_1, w_2 που υπολογίσαμε ως στήλες είναι ο ταυτικός. Αυτό είναι προφανές, αφού ξεκινήσαμε να εκτελούμε την μέθοδο εύρεσης του πίνακα Jordan σε ένα πίνακα που ήταν ήδη σε αυτή την μορφή.

2. **Τρία γραμμικά ανεξάρτητα ιδιοδιανύσματα** Ας προσπαθήσουμε να υπολογίσουμε την κανονική μορφή Jordan του πίνακα

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο $-(x-3)^2(x-5)$ το οποίο έχει τις ιδιοτιμές $\lambda = 3$ και $\lambda = 5$. Υπολογίζουμε τους ιδιόχωρους E_3, E_5 λύνοντας τα ομογενή συστήματα:

$$(A - 3I_3)x = 0, \quad (A - 5I_3)x = 0$$

και έχουμε ότι

$$E_3 = \langle (0, 1, 0)^t, (-1, 0, 1)^t \rangle, E_5 = \langle (1, 2, 1)^t \rangle.$$

Παρατηρούμε ότι έχουμε βρει $3 = \dim \mathbb{R}^3$ γραμμικά ανεξάρτητα διανύσματα. Ο πίνακας είναι διαγωνοποιήσιμος. Θέτουμε

$$P = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

και υπολογίζουμε

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -1 & 1 & -1 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \quad P^{-1}AP = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Ο πίνακας είναι διαγωνοποιήσιμος και τα Jordan blocks είναι οι 1×1 πίνακες $[5], [3], [3]$.

3. **Δύο γραμμικά ανεξάρτητα ιδιοδιανύσματα** Θεωρούμε τον πίνακα

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

το οποίο έχει χαρακτηριστικό πολυώνυμο $-(x-1)^3$. Υπολογίζουμε μια βάση του ιδιόχωρου E_1 της ιδιοτιμής $\lambda = 1$ ως την λύση του συστήματος

$$(A - \mathbb{I}_3)x = \mathbf{0} \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

$$E_1 = \langle (1, 0, 0)^t, (0, 1, -1)^t \rangle.$$

Προσπαθούμε να βρούμε ένα διάνυσμα w_1 που να ικανοποιεί το σύστημα

$$(A - \mathbb{I}_3)w_1 = (0, 1, -1)^t \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Το σύστημα αυτό όμως είναι αδύνατο και η διαδικασία στο βήμα 3 της μεθόδου **5** σταματάει. Προσπαθούμε τώρα να βρούμε ένα διάνυσμα w_1 που να ικανοποιεί το σύστημα

$$(A - \mathbb{I}_3)w_1 = (1, 0, 0)^t \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Το σύστημα αυτό έχει λύσεις (παρατηρήστε ότι το αντίστοιχο ομογενές το έχουμε ήδη λύσει)

$$x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Διαλέγουμε στον χώρο λύσεων την απλούστερη λύση, δηλαδή την $w_1 = (0, 0, 1)^t$. Έχουμε ήδη βρει τρία γραμμικά ανεξάρτητα διανύσματα, τα $(1, 0, 0)^t, (0, 1, -1)^t, (0, 0, 1)^t$ οπότε μπορούμε να σταματήσουμε εδώ. Αν προσπαθούσαμε να συνεχίσουμε θα ψάχναμε για μια λύση στο σύστημα

$$(A - \mathbb{I}_3)w_2 = (0, 0, 1)^t \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

θα βλέπαμε ότι το σύστημα είναι αδύνατο. Μπορούμε τώρα να θεωρήσουμε τον πίνακα

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

και να επαληθεύσουμε ότι

$$P^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

4. **Ένα γραμμικά ανεξάρτητο ιδιοδιάνυσμα** Σε αυτό το παράδειγμα θεωρούμε τον πίνακα

$$A = \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{pmatrix}$$

του οποίου θέλουμε να υπολογίσουμε την κανονική μορφή Jordan του. Υπολογίζουμε ότι το χαρακτηριστικό του πολυώνυμο είναι το $-(\lambda + 1)^3$, συνεπώς έχουμε μοναδική ιδιοτιμή $\lambda = -1$ και τον ιδιόχωρο $E_{-1} = \{x : \mathbb{R}^3 : (A + \mathbb{I}_3)x = 0\}$, ο οποίος έχει διάσταση ένα και παράγεται από το $v = (1, 0, 0)^t$. Αναζητούμε ένα διάνυσμα

$$(A + \mathbb{I}_3)w_1 = v \Leftrightarrow \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Παρατηρούμε ότι ο χώρος λύσεων είναι το $x(1, 0, 0)^t + (0, -1, 0)^t$, $x \in \mathbb{R}$ και διαλέγουμε την απλούστερη λύση $w_1 = (0, -1, 0)^t$. Συνεχίζουμε, αναζητώντας τις λύσεις του συστήματος

$$(A + \mathbb{I}_3)w_1 = v \Leftrightarrow \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

το οποίο έχει λύσεις $x(1, 0, 0)^t + (0, 0, 1/2)^t$, $x \in \mathbb{R}$. Και πάλι διαλέγουμε την λύση $w_2 = (0, 0, 1/2)^t$. Τα v, w_1, w_2 είναι γραμμικά ανεξάρτητα και έχουμε τελειώσει. Αν επιχειρούσαμε να βρούμε μια λύση $(A + \mathbb{I}_3)w_3 = w_2$, θα διαπισώναμε ότι το τελευταίο σύστημα είναι αδύνατο. Διαλέγουμε τον πίνακα P

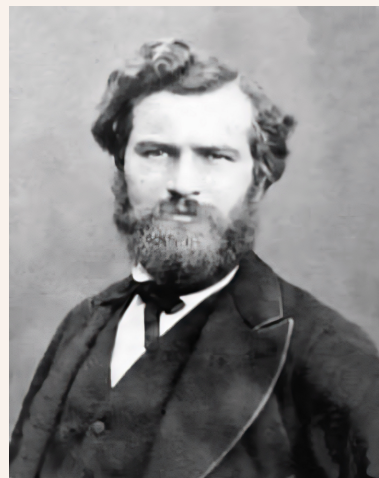
$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

και επαληθεύουμε

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Ο **Marie Ennemond Camille Jordan** (5 Ιανουαρίου 1838 - 22 Ιανουαρίου 1922) ήταν Γάλλος μαθηματικός με σημαντική συνεισφορά στην θεωρία ομάδων και ιδιαίτερα γνωστός για το σύγγραμμά του *Cours d'analyse*.

Άφησε το όνομα του σε μια σειρά από αποτελέσματα, όπως το θεώρημα της κλειστής καμπύλης του Jordan στην τοπολογία ή το θεώρημα Jordan-Hölder στις συνθετικές σειρές στην θεωρία ομάδων την έννοια των οποίων εισήγαγε όπως και έθεσε το πρόβλημα της ταξινόμησης των επιλύσιμων ομάδων. Την εργασία του στην θεωρία των ομάδων δημοσίευσε το 1870 στο πρώτο ιστορικά βιβλίο θεωρίας ομάδων: *Traité des substitutions et des équations algébriques*, το οποίο περιείχε και μια περιεκτική μελέτη της θεωρίας Galois. Επίσης είχε συνεισφορά στην μελέτη των ομάδων πινάκων τόσο πάνω από σώματα χαρακτηριστικής μηδέν όσο και πάνω από σώματα θετικής χαρακτηριστικής.



Θεώρημα VI.3.24. Για κάθε πίνακα $A \in \mathbb{C}^{n,n}$ υπάρχει μια ακολουθία διαγωνοποιήσιμων πινάκων $A_n \rightarrow A$.

Απόδειξη. Κάθε πίνακας A είναι όμοιος με ένα άνω τριγωνικό πίνακα B ο οποίος έχει ιδιοτιμές $\lambda_1, \dots, \lambda_k$, δηλαδή υπάρχει αντιστρέψιμος Q ώστε $Q^{-1}AQ = B$. Αν ο πίνακας έχει διαφορετικές ιδιοτιμές είναι διαγωνοποιήσιμος. Αν ο πίνακας έχει κάποιες ιδιοτιμές ίδιες, τότε μπορούμε να βρούμε μια ακολουθία από διαγώνιους πίνακες D_n με $\lim_{n \rightarrow \infty} D_n = \mathbf{0}_{n,n}$ ώστε οι πίνακες $B + D_n$ να έχουν διαφορετικές ιδιοτιμές για κάθε $n \in \mathbb{N}$ και συνεπώς είναι διαγωνοποιήσιμοι. Ισχύει ότι

$$\lim_{n \rightarrow \infty} Q(B + D_n)Q^{-1} = A,$$

δηλαδή το ζητούμενο. □

Παράδειγμα VI.3.25. Οι πίνακες

$$A_n = \begin{pmatrix} 1 + \frac{1}{n} & 1 & 0 \\ 0 & 1 + \frac{2}{n} & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = A.$$

Οι πίνακες A_n έχουν δύο διαφορετικές ιδιοτιμές και είναι διαγωνοποιήσιμοι, ενώ ο πίνακας A δεν είναι. Ο πίνακας D_n στην απόδειξη του θεωρήματος VI.3.24 είναι ο $\text{diag}(1/n, 2/n, 0)$. Είναι ενδιαφέρον ότι μια βάση από ιδιοδιανύσματα είναι οι στήλες του πίνακα

$$Q = \begin{pmatrix} \frac{n^2}{2} & 1 & n \\ -\frac{n}{2} & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

οι οποίες δεν έχουν νόημα στο όριο $n \rightarrow \infty$ εκτός από το ιδιοδιάνυσμα $(1, 0, 0)^t$.

Παρατήρηση VI.3.26 (Αναλυτική απόδειξη του θεωρήματος Caley-Hamilton). Παρατηρούμε ότι το Caley-Hamilton ισχύει για διαγωνοποιήσιμους πίνακες. Πράγματι, το χαρακτηριστικό πολυώνυμο ενός διαγωνοποιήσιμου πίνακα $A = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^{-1}$ είναι

$$\text{Ch}_A(x) = \prod_{i=1}^n (x - \lambda_i)$$

$$\text{Ch}_A(A) = \prod_{i=1}^n (A - \lambda_i \mathbb{I}_n) = Q \prod_{i=1}^n \text{diag}(\lambda_1 - \lambda_i, \dots, \lambda_n - \lambda_i) Q^{-1} = \mathbf{0}_{\mathbb{F}^{n,n}}.$$

Για την τελευταία ισότητα παρατηρήστε ότι έχουμε ένα γινόμενο $n \times n$ διαγώνιων πινάκων, όπου ο κάθε ένας έχει τουλάχιστον στην i θέση της διαγωνίου ένα μηδενικό στοιχείο.

Για ένα μη διαγωνοποιήσιμο πίνακα A , αν το $\mathbb{F} = \mathbb{R}$ ή $\mathbb{F} = \mathbb{C}$ μπορούμε να βρούμε μια ακολουθία διαγωνοποιήσιμων πινάκων $A_n \rightarrow A$. Το χαρακτηριστικό πολυώνυμο είναι μια συνεχής συνάρτηση $\mathbb{F}^{n,n} \rightarrow \mathbb{F}[x]$ οπότε

$$\mathbf{0}_{n,n} = \lim_{n \rightarrow \infty} \text{Ch}_{A_n}(A_n) = \text{Ch}_A(\lim_{n \rightarrow \infty} A_n) = \text{Ch}_A(A),$$

από όπου έχουμε το ζητούμενο.

Ασκήσεις

Άσκηση VI.17 Θεωρούμε ένα 6×6 πίνακα με ελάχιστο πολυώνυμο

$$m(x) = (x-1)^2(x-2)$$

και χαρακτηριστικό πολυώνυμο

$$f(x) = (x-1)^4(x-2)^2.$$

Να βρεθεί η πιθανή μορφή Jordan του πίνακα.

Άσκηση VI.18 Να βρεθεί η μορφή Jordan του πίνακα

$$\begin{pmatrix} 4 & -4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 & -3 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 8 & -3 & 8 \end{pmatrix}$$

VI.4 Εφαρμογές στις αναδρομικές ακολουθίες

Ας υποθέσουμε ότι έχουμε την αναδρομική ακολουθία

$$u_0 = 1, u_n = Au_0, A \in \mathbb{F}.$$

Είναι σαφές ότι ο γενικός όρος της ακολουθίας δίνεται από τον τύπο

$$u_n = A^n u_0 = A^n.$$

Την διαδικασία αυτή θέλουμε να την γενικεύσουμε. Θα ξεκινήσουμε από ένα κλασικό πρόβλημα.

Η ακολουθία Fibonacci είναι μια αναδρομική ακολουθία η οποία ορίζεται ως εξής: θέτουμε $F_0 = F_1 = 1$ και

$$F_n = F_{n-1} + F_{n-2}$$

δηλαδή κάθε όρος της είναι το άθροισμα των δύο προηγούμενων. Παρατηρούμε ότι η παραπάνω αναδρομική σχέση μπορεί να εφραστεί με την βοήθεια του πινάκων:

$$\begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-2} \\ F_{n-1} \end{pmatrix} \quad (\text{VI.17})$$

Αν λοιπόν θέσουμε

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad u_n = \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix}$$

τότε η ακολουθία Fibonacci δίνεται από τον τύπο

$$u_n = Au_{n-1}, \quad u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

και συνεπώς ο γενικός όρος της δίνεται από τον τύπο

$$u_{n+1} = \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n u_1.$$

Συνεπώς χρειαζόμαστε έναν αποτελεσματικό τρόπο να υπολογίζουμε δυνάμεις του πίνακα A . Στο παράδειγμα VI.1.21 δείξαμε ότι

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = Q \begin{pmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{pmatrix} Q^{-1}, \text{ όπου } Q = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & \frac{-1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}.$$

Συνεπώς,

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n u_0 = Q \begin{pmatrix} \left(\frac{1-\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1+\sqrt{5}}{2}\right)^n \end{pmatrix} Q^{-1} u_0 = \begin{pmatrix} \frac{2^{-n}((1+\sqrt{5})^n - (1-\sqrt{5})^n)}{\sqrt{5}} \\ \frac{2^{-2n-1}((\sqrt{5}-1)(2-2\sqrt{5})^n + 2^n(1+\sqrt{5})^{n+1})}{\sqrt{5}} \end{pmatrix}$$

Από όπου προκύπτει ότι

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

Ο παραπάνω τύπος ανακατώνει ρητές δυνάμεις του άρρητου αριθμού $\sqrt{5}$ οι οποίες με κάποιο μαγικό τρόπο αλληλοαναιρούνται για να δώσουν ως τελικό αποτέλεσμα έναν ακέραιο.

Ασκήσεις

Άσκηση VI.19

- Είναι διαγωνοποιήσιμος ο πίνακας

$$A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix};$$

- Θεωρούμε δύο ακολουθίες (u_n) και (v_n) που ορίζονται ως εξής: $u_0 = v_0 = 1$, $u_{n+1} = 6u_n - 2v_n$, $v_{n+1} = -2u_n + 9v_n$. Να υπολογιστούν τα u_n, v_n συναρτήσει του n .

VI.4.1 Ρητή Κανονική μορφή

Στην περίπτωση που ο διανυσματικός χώρος είναι ορισμένος υπέρ ενός αλγεβρικά κλειστού σώματος $\mathbb{F}$ η κανονική μορφή Jordan είναι το καλύτερο που μπορούμε να κάνουμε για να περιγράψουμε τις κλάσεις ομοιότητας πινάκων. Στην παράγραφο αυτή θα εξετάσουμε την ρητή κανονική μορφή η οποία δεν προϋποθέτει να είναι το σώμα $\mathbb{F}$ αλγεβρικά κλειστό. Στην πορεία θα επανεξετάσουμε από μια διαφορετική οπτική την θεωρία των πολυωνύμων γραμμικών συναρτήσεων.

Modules

Δίνεται μια γραμμική συνάρτηση $L: V \rightarrow V$. Μπορούμε να ορίσουμε μια συνάρτηση

$$\begin{aligned} \mathbb{F}[x] \times V &\longrightarrow V \\ (f, v) &\longmapsto f(L)v =: fv \end{aligned} \tag{VI.18}$$

Την παραπάνω συνάρτηση την θεωρούμε ως ένα «πολλαπλασιασμό» πολυωνύμων με στοιχεία του διανυσματικού χώρου V . Ο πολλαπλασιασμός αυτός φυσικά εξαρτάται από την γραμμική συνάρτηση V και ικανοποιεί τις ιδιότητες:

- $(f + g)v = (f + g)(L)v = f(L)v + g(L)v = fv + gv$ για κάθε $f, g \in \mathbb{F}[x]$ και κάθε $v \in V$.

2. $f(v + w) = f(L)(v + w) = f(L)v + f(L)w = fv + fw$, για κάθε $f \in \mathbb{F}[x]$ και κάθε $v, w \in V$
3. $(fg)v = fg(L)v = f(L)g(L)v = f(gv)$, για κάθε $f, g \in \mathbb{F}[x]$ και κάθε $v \in V$.
4. $1v = \text{Id}_V v = v$.

Η παρακάτω κατασκευή μας δίνει την ανάγκη να ορίσουμε την έννοια του module (το οποίο στην Ελληνική βιβλιογραφία αναφέρεται συχνά ως «πρότυπο»)

Ορισμός VI.4.1. Μία αβελιανή ομάδα $(V, +)$ θα είναι ένα R -module πάνω από τον αντιμεταθετικό δακτύλιο R , ο οποίος έχει μονάδα 1_R , αν υπάρχει μια συνάρτηση

$$\begin{aligned} R \times V &\longrightarrow V \\ (f, v) &\longmapsto f(L)v =: fv \end{aligned}$$

για την οποία να ισχύουν οι ιδιότητες:

1. $(f + g)v = fv + gv$ για κάθε $f, g \in R$ και κάθε $v \in V$.
2. $f(v + w) = fv + fw$, για κάθε $f \in R$ και κάθε $v, w \in V$
3. $(fg)v = f(gv)$, για κάθε $f, g \in R$ και κάθε $v \in V$.
4. $1_R v = v$.

Παρατήρηση VI.4.2. Κάθε αβελιανή ομάδα V αποκτά με φυσιολογικό τρόπο δομή $\mathbb{Z}$ -module θέτοντας

$$nv = \begin{cases} \underbrace{v + \dots + v}_{n\text{-φορες}} & \text{αν } n > 0 \\ 0 & \text{αν } n = 0 \\ \underbrace{(-v) + \dots + (-v)}_{-n\text{-φορες}} & \text{αν } n < 0 \end{cases}$$

Στην πραγματικότητα η έννοια της αβελιανής ομάδας και του $\mathbb{Z}$ -module ταυτίζονται.

Παρατήρηση VI.4.3. Η έννοια του R -module γενικεύει την έννοια του διανυσματικού χώρου. Πράγματι ένας διανυσματικός χώρος πάνω από το σώμα $\mathbb{F}$ είναι το ίδιο με ένα $\mathbb{F}$ -module, όπου έχουμε απαιτήσει ότι ο δακτύλιος R να είναι το σώμα $\mathbb{F}$.

Όμως υπάρχουν και σοβαρές διαφορές. Πράγματι, σε ένα R -module μπορεί να έχουμε ότι $fv = 0$ με $f \in R$, $f \neq 0$ και $v \in V$, $v \neq 0$. Σκεφτείτε το $5(1 \bmod 5) = 0$ στην ομάδα $V = \mathbb{Z}/5\mathbb{Z}$. Για τις ανάγκες της γραμμικής άλγεβρας έχουμε ότι αν $m \neq 0$ είναι το ελάχιστο πολυώνυμο του L , τότε $mv = m(L)v = 0_V$ για κάθε $v \in V$.

Ορισμός VI.4.4. Ένα υποσύνολο W ενός R -module θα λέγεται υποmodule αν και μόνο αν οι περιορισμοί των πράξεων της πρόσθεσης και του και του εξωτερικού πολλαπλασιασμού $R \times V \rightarrow V$ στο W , το εφοδιάζουν με την δομή ενός R -module.

Παρατήρηση VI.4.5. Αυτό που χρειάζεται να ελεγχθεί είναι κατά πόσο οι πράξεις είναι κλειστές στο W , δηλαδή

- $x - y \in W$ για κάθε $x, y \in W$
- $rx \in W$ για κάθε $r \in R$ και $x \in W$.

Παρατήρηση VI.4.6. Κάθε υποσύνολο I του R είναι R -module αν και μόνο αν είναι ιδεώδες του R .

Παρατήρηση VI.4.7. Στην περίπτωση που έχουμε την δράση του $\mathbb{F}[x]$ στον διανυσματικό χώρο V , μέσω της γραμμικής συνάρτησης $L : V \rightarrow V$, έχουμε ότι ένας υπόχωρος $W \subset V$ είναι R -υποmodule αν και μόνο αν είναι L -αναλλοίωτος.

Ορισμός VI.4.8. Μία συνάρτηση $f : V_1 \rightarrow V_2$ ανάμεσα σε R -modules θα λέγεται μορφισμός από R -modules αν διατηρεί τις πράξεις, δηλαδή αν

- $f(v_1 + v_2) = f(v_1) + f(v_2)$ για κάθε $v_1, v_2 \in V_1$
- $f(rv) = rf(v)$ για κάθε $v \in V_1$ και κάθε $r \in R$.

Ένας μορφισμός από R -modules που είναι 1-1 θα λέγεται μονομορφισμός από R -modules, ενώ ένας μορφισμός από R -modules που είναι επί θα λέγεται επιμορφισμός από R -modules. Τέλος ένας μορφισμός από R -modules που είναι και 1-1 και επί θα λέγεται ισομορφισμός από R -modules.

Παρατήρηση VI.4.9. Οι παραπάνω ορισμοί γενικεύουν την έννοια της γραμμικής συνάρτησης που είναι ένας μορφισμός από $\mathbb{F}$ -modules.

Ορισμός VI.4.10. Ένα R -module V θα λέγεται ελεύθερο τάξης n , αν υπάρχουν στοιχεία $v_1, \dots, v_n \in V$, ώστε κάθε στοιχείο $v \in V$ να γράφεται με μοναδικό τρόπο ως

$$v = r_1 v_1 + r_2 v_2 + \dots + r_n v_n, \text{ για } r_1, \dots, r_n \in R.$$

Παρατήρηση VI.4.11. Ένα ελεύθερο R -module τάξης n είναι ισόμορφο με το R -module, R^n το οποίο αποτελείται από το σύνολο των n -άδων

$$R^n = (r_1, r_2, \dots, r_n)$$

πρόσθεση κατά συντεταγμένες και εξωτερικό πολλαπλασιασμό $r(r_1, r_2, \dots, r_n) = (rr_1, rr_2, \dots, rr_n)$. Πράγματι, αρκεί να θεωρήσουμε την συνάρτηση

$$\begin{aligned} R^n &\longrightarrow V \\ (r_1, \dots, r_n) &\longmapsto v = \sum_{i=1}^n r_i e_i \end{aligned}$$

και να τσεκάρουμε ότι είναι ισομορφισμός.

Παρατήρηση VI.4.12. Δύο ισόμορφα ελεύθερα R -modules, με R -αντιμεταθετικός δακτύλιος έχουν την ίδια τάξη. Δηλαδή το R^n δεν μπορεί να είναι ισόμορφο με το R^m για $n \neq m$. Ισχυριζόμαστε ότι κάθε αντιμεταθετικός δακτύλιος έχει ένα τουλάχιστον μέγιστο ιδεώδες, έστω $\mathfrak{m}$. Τότε έχουμε ότι $R^n/\mathfrak{m}R^n \cong k^n$ ενώ $R^m/\mathfrak{m}R^m = k^m$, όπου k είναι το σώμα $R/\mathfrak{m}$. Ένας ισομορφισμός ανάμεσα στο R^n και το R^m θα επάγει ένα ισομορφισμό ανάμεσα στους διανυσματικούς χώρους k^n και k^m για $n \neq m$, άτοπο.

Για να αποδείξουμε τώρα ότι ένας αντιμεταθετικός δακτύλιος R έχει ένα μέγιστο ιδεώδες, θα χρησιμοποιήσουμε το λήμμα του Zorn, [I.4.20](#). Θα αποδείξουμε μάλιστα ότι κάθε ιδεώδες $\neq R$ περιέχεται σε μέγιστο ιδεώδες του R . Θεωρούμε το σύνολο Σ όλων των ιδεωδών του R τα οποία είναι διάφορα του R και περιέχουν το R . Το σύνολο R είναι επαγωγικό ως προς την διάταξη του

εγκλεισμού, δηλαδή κάθε ολικά διατεταγμένο υποσύνολο του έχει άνω φράγμα που ανήκει στο Σ . Πράγματι, αν $\{I_\nu\}_{\nu \in A}$ είναι ένα ολικά διατεταγμένο υποσύνολο του Σ τότε το $\cup_{\nu \in A} I_\nu$ είναι ένα ιδεώδες του R , το οποίο είναι διάφορο του R και αποτελεί ένα άνω φράγμα του $\{I_\nu\}_{\nu \in A}$. Έστω ένα μέγιστο στοιχείο m του Σ , η ύπαρξη του οποίου εξασφαλίζεται από το λήμμα του Zorn. Αυτό είναι μέγιστο ιδεώδες, γιατί διαφορετικά θα υπήρχε $m \subset J \subsetneq R$, και είναι σαφές ότι $J \in \Sigma$, το οποίο είναι άτοπο με βάση τον ορισμό του m .

Παρατήρηση VI.4.13. Κάθε πεπερασμένα παραγόμενος διανυσματικός χώρος είναι ελεύθερο $\mathbb{F}$ -module. Αυτό σημαίνει ότι έχει μια βάση. Στην περίπτωση των R -modules το αντίστοιχο θεώρημα δεν είναι σωστό. Έτσι το πεπερασμένο $\mathbb{Z}$ -module $\mathbb{Z}/5\mathbb{Z}$ δεν είναι ελεύθερο, αφού δεν είναι ισόμορφο με το $\mathbb{Z}^n$ που έχει άπειρα το πλήθος στοιχεία.

Ορισμός VI.4.14. Ένα R -module V θα λέγεται *torsion module* αν για κάθε $v \in V$ υπάρχει $r \in R$, ώστε $rv = 0$.

Παράδειγμα VI.4.15. Κάθε πεπερασμένης διάστασης $\mathbb{F}$ -διανυσματικός χώρος, εφοδιασμένος με την δομή ενός $\mathbb{F}[x]$ -module μέσω μιας γραμμικής συνάρτησης $L : V \rightarrow V$ είναι ένα *torsion module*, αφού για το ελάχιστο πολυώνυμο m_L του L , έχουμε $m_L v = 0_V$.

Από εδώ και στο εξής ο δακτύλιος R θα είναι μια περιοχή κυρίων ιδεωδών. Το βασικό παράδειγμα το οποίο έχουμε στο μυαλό μας για τις ανάγκες της γραμμικής άλγεβρας είναι ο $R = \mathbb{F}[x]$.

Θεωρούμε το R -module V , και έστω $v \in V$. Θεωρούμε την συνάρτηση $\phi : R \rightarrow Rv$ η οποία στέλνει το $r \in R$ στο $rv \in Rv$. Θεωρούμε τον πυρήνα της συνάρτησης ϕ ο οποίος είναι ένα κύριο ιδεώδες το οποίο παράγεται από ένα στοιχείο $m_v \in R$. Το στοιχείο m_v θα το λέμε περίοδο του v in V . Ένα στοιχείο m θα λέγεται ένας εκθέτης για το V (αντίστοιχα για το $v \in V$) αν $mV = 0$ (αντίστοιχα $mv = 0$). Για παράδειγμα στην δράση του $\mathbb{F}[x]$ στον διανυσματικό χώρο V , μέσω της γραμμικής συνάρτησης $L : V \rightarrow V$, το ελάχιστο πολυώνυμο είναι ένας εκθέτης για το V . Για ένα $m \in R$ ορίζουμε το σύνολο

$$V_m = \{v \in V : mv = 0\},$$

δηλαδή τα στοιχεία του V που έχουν εκθέτη m . Ένα R -module Θ λέγεται *κυκλικό* αν είναι ισόμορφο με το $R/\langle a \rangle$, για κάποιο $a \in R$, $a \neq 0$. Το a θα το λέμε τάξη του κυκλικού module. Για ένα ανάγωγο στοιχείο $p \in R$, ορίζουμε

$$V(p) = \{v \in V : pv = 0\},$$

δηλαδή το υποmodule του V των στοιχείων του V που έχουν εκθέτη p .

Θεώρημα VI.4.16. Έστω V ένα μη-μηδενικό πεπερασμένο *torsion* R -module. Τότε

$$V = \bigoplus_p V(p), \quad (\text{VI.19})$$

όπου το άθροισμα λαμβάνεται πάνω σε όλα τα ανάγωγα p με $V(p) \neq 0$. Κάθε $V(p)$ με την σειρά του γράφεται ως

$$V(p) = R/\langle p^{v_1} \rangle \oplus \cdots \oplus R/\langle p^{v_s} \rangle, \quad (\text{VI.20})$$

με $1 \leq v_1 \leq \cdots \leq v_s$, όπου οι ακολουθία $v_1, \dots, v_s$ είναι μονοσήμαντα ορισμένη.

Απόδειξη. Έστω m ένας εκθέτης του V τον οποίο αναλύουμε σε γινόμενο πρώτων μεταξύ τους ανάγωγων πολυωνύμων. Όπως ακριβώς στην απόδειξη της πρότασης VI.3.6 αποδεικνύουμε την αλήθεια της ανάλυσης στην εξίσωση (VI.19).

Θα προχωρήσουμε τώρα στην απόδειξη της (VI.20). Τα στοιχεία $v_1, \dots, v_n \in V$ θα λέγονται εξαρτημένα αν οποιεδήποτε έχουμε μια σχέση

$$r_1 v_1 + \dots + r_n v_n = 0$$

με $r_1, \dots, r_n \in R$ ισχύει ότι $r_i v_i = 0$, για κάθε $1 \leq i \leq n$. Η συνθήκη αυτή είναι ασθενέστερη της γραμμικής ανεξαρτησίας. Είναι σαφές ότι τα στοιχεία $v_1, \dots, v_n$ είναι ανεξάρτητα αν και μόνο αν έχουμε την παρακάτω ανάλυση από R -modules:

$$\langle v_1, \dots, v_n \rangle = \langle v_1 \rangle \oplus \dots \oplus \langle v_n \rangle,$$

όπου τα $\langle v_i \rangle = Rv_i \cong R/m_{v_i} R$, είναι κυκλικά modules.

Θεωρούμε ένα torsion module εκθέτη p^r , $r \geq 1$ για κάποιο ανάγωγο p . Έστω ένα στοιχείο v με περίοδο p^r . Θεωρούμε το πηλίκο $\bar{V} = V/\langle v \rangle$. Έστω $\bar{w}_1, \dots, \bar{w}_m$ ανεξάρτητα στοιχεία του $\bar{V}$. Θα δείξουμε ότι για κάθε ένα $\bar{w}_i$ υπάρχει ένας αντιπρόσωπος w_i της κλάσης modulo $\langle v \rangle$ ώστε ο εκθέτης του $\bar{w}_i$ να είναι ίδιος με τον εκθέτη του w_i .

Πράγματι, έστω $\bar{w}$ να έχει περίοδο p^n για κάποιο $n \geq 1$. Έστω w ένας αντιπρόσωπος της κλάσης του $\bar{w}$. Ισχύει ότι $p^n w \in \langle v \rangle$, συνεπώς $p^n w = p^s c v$, με $c \in R$, $(c, p) = 1$, $0 \leq s \leq r$. Αν $s = r$ τότε $p^n w = 0$ και το w , έχει τον ίδιο εκθέτη όπως και το $\bar{w}$. Αν το $s < r$, τότε το $p^s c v$ έχει περίοδο p^{r-s} και συνεπώς το w έχει περίοδο p^{n+r-s} . Είναι σαφές ότι $n+r-s \leq r$, αφού το p^r είναι εκθέτης του V . Συνεπώς $n \leq s$ και το στοιχείο $w - p^{s-n} c v$ είναι ένας αντιπρόσωπος του $\bar{w}$ με εκθέτη n .

Διαλέγουμε αντιπροσώπους w_i των κλάσεων $\bar{w}_i$ με την ίδια περίοδο. Θα αποδείξουμε ότι τα $v, w_1, \dots, w_m$ είναι ανεξάρτητα. Έστω $r, r_1, \dots, r_m \in R$ ώστε

$$rv + r_1 w_1 + \dots + r_m w_m = 0. \quad (\text{VI.21})$$

Είναι σαφές ότι

$$r_1 \bar{w}_1 + \dots + r_m \bar{w}_m = 0$$

Από την υπόθεση της ανεξαρτησίας των $\bar{w}_1, \dots, \bar{w}_m$ έχουμε ότι $r_i \bar{w}_i = 0$ για όλα τα i . Αν p^{v_i} είναι η περίοδος του $\bar{w}_i$ τότε p^{v_i} διαιρεί το r_i . Τότε όμως $r_i w_i = 0$ για κάθε i . Επιστρέφοντας στην εξίσωση (VI.21) έχουμε ότι και $rv = 0$.

Διαλέγουμε το $v \in V(p)$ με τέτοιο τρόπο ώστε η περίοδος του p^r να έχει μέγιστο δυνατό r . Παρατηρούμε ότι το V_p , το οποίο είναι το σύνολο των στοιχείων με περίοδο p , γίνεται ένας $k = R/pR$ -διανυσματικός χώρος. Έχουμε ότι $\dim_k \bar{V}_p < \dim_k V_p$, αφού μπορούμε στο $\langle v \rangle$ να βρούμε στοιχείο περιόδου p το οποίο να είναι γραμμικά ανεξάρτητο από μια βάση του $\bar{V}_p$.

Θα ολοκληρώσουμε την απόδειξη του θεωρήματος με επαγωγή. Αν το $\bar{V}$ δεν είναι μηδενικό, τότε υπάρχουν στοιχεία $\bar{v}_2, \dots, \bar{v}_s$ τα οποία έχουν αντίστοιχες περιόδους $p^{r_2}, \dots, p^{r_s}$ ώστε $r_2 \geq \dots \geq r_s$. Διαλέγουμε αντιπροσώπους $v_2, \dots, v_s$ στο V με τις ίδιες περιόδους ώστε τα $v_1, v_2, \dots, v_s$ να είναι ανεξάρτητα. Αφού το p^{r_1} έχει διαλεχτεί με τέτοιο τρόπο ώστε r_1 να είναι μέγιστο, η ανισότητα $r_1 \geq r_2 \geq \dots$, είναι σαφής.

Η μοναδικότητα θα αποδειχθεί ως συνέπεια του γενικότερου θεωρήματος VI.4.17. $\square$

Θεώρημα VI.4.17. Έστω V ένα πεπερασμένο παραγόμενο μη-μηδενικό torsion R -module. Τότε το V είναι ισόμορφο με ένα ευθύ άθροισμα

$$V = R/\langle q_1 \rangle \oplus \dots \oplus R/\langle q_r \rangle, \quad (\text{VI.22})$$

όπου $q_1, \dots, q_r$ είναι μη-μηδενικά στοιχεία του R , και $q_1 \mid q_2 \mid \dots \mid q_r$. Η ακολουθία των ιδεωδών $\langle q_1 \rangle, \dots, \langle q_r \rangle$, είναι μονοσήμαντα ορισμένη από τις παραπάνω συνθήκες.

Απόδειξη. Με χρήση του θεωρήματος VI.4.16 διασπάμε το V ως

$$V = V(p_1) \oplus \cdots \oplus V(p_l),$$

όπου κάθε $V(p_i)$ διασπάται με την σειρά του ως ευθύ άθροισμα κυκλικών υποmodules περιόδων $p_i^{r_{ij}}$. Κρατάμε την πληροφορία των εκθετών στον παρακάτω πίνακα:

$$\begin{array}{c|c} V(p_1) & r_{11} \leq r_{12} \leq \cdots \\ V(p_2) & r_{21} \leq r_{22} \leq \cdots \\ \vdots & \vdots \\ V(p_l) & r_{l1} \leq r_{l2} \leq \cdots \end{array}$$

Θέτουμε τα $q_1, q_2, \dots, q_r$ να είναι τα γινόμενα κατά μήκος των παραπάνω στηλών, δηλαδή

$$q_1 = p_1^{r_{11}} p_2^{r_{21}} \cdots p_l^{r_{l1}}, \quad q_2 = p_1^{r_{12}} p_2^{r_{22}} \cdots p_l^{r_{l2}}, \dots$$

Από το θεώρημα του Κινέζου **να γράφει στους δακτυλίους και να δοθεί αναφορά** το $R/\langle q_1 \rangle$ είναι ισόμορφο με το ευθύ άθροισμα των modules κατά μήκος της πρώτης στήλης δηλαδή

$$R/\langle q_1 \rangle = R/\langle p_1^{r_{11}} \rangle \oplus R/\langle p_2^{r_{21}} \rangle \oplus \cdots \oplus R/\langle p_l^{r_{l1}} \rangle$$

ομοίως και για τα $R/\langle q_2 \rangle, \dots, R/\langle q_r \rangle$. Με αυτό τον τρόπο αποδεικνύεται η εξίσωση VI.22, ενώ η συνθήκη $q_1 \mid q_2 \mid \cdots \mid q_r$ είναι προφανής από τις ανισότητες των εκθετών.

Για την μοναδικότητα τώρα. Θεωρούμε ένα ανάγωγο στοιχείο p και θεωρούμε την περίπτωση που $V' = R/\langle pb \rangle$ για κάποιο $b \in R$, $b \neq 0$. Το V'_p είναι το υποmodule $bR/\langle pb \rangle$, όπως προκύπτει από την μοναδική παραγοντοποίηση του R . Θεωρούμε την συνάρτηση $V' \rightarrow pV'$ η οποία είναι επί και έχει πυρήνα V'_p . Συνεπώς

$$pV' = V'/V'_p = \frac{R/\langle pb \rangle}{bR/\langle pb \rangle} \cong R/bR.$$

Θεωρούμε τώρα το V να εκφράζεται όπως το θεώρημα VI.4.17 ως ένα ευθύ άθροισμα από r προσθετέους. Το τυχαίο στοιχείο $v \in V$ γράφεται

$$v = v_1 + v_2 + \cdots + v_r$$

και είναι στοιχείο του V_p αν και μόνο αν $pv_i = 0$ για όλα τα i , $1 \leq i \leq r$. Δηλαδή το V_p είναι το ευθύ άθροισμα των πυρήνων του πολλαπλασιασμού με p σε κάθε προσθετέο. Όμως το V_p , αφού ο πολλαπλασιασμός με p είναι μηδενικός, είναι ένας διανυσματικός χώρος πάνω από το $k = R/\langle p \rangle$ και η διάστασή του είναι ίση με το πλήθος των $R/\langle q_i \rangle$ που το $p \mid q_i$.

Αν το p είναι ανάγωγο που διαιρεί το q_1 και συνεπώς όλα τα επόμενα q_i για $1 \leq i \leq r$. Υποθέτουμε ότι V έχει και μια δεύτερη ανάλυση

$$V = R/\langle q'_1 \rangle \oplus \cdots \oplus R/\langle q'_s \rangle,$$

η οποία ικανοποιεί τις συνθήκες του θεωρήματος. Τότε θα πρέπει το p να διαιρεί τουλάχιστον r από τα στοιχεία q'_j και συνεπώς $r \leq s$. Για λόγους συμμετρίας έχουμε ότι $r = s$ και ότι το p διαιρεί όλα τα q'_j .

Θεωρούμε το module pV . Αν γράψουμε $q_i = pb_i$ τότε

$$pV \cong R/\langle b_1 \rangle \oplus \cdots \oplus R/\langle b_r \rangle,$$

και $b_1 \mid \cdots \mid b_r$. Κάποια από τα b_i μπορεί να είναι μονάδες. Αν όχι συνεχίζουμε να βγάζουμε έξω και άλλους πρώτους διαιρέτες του b_1 και συνεπώς όλως των b_i . Κάποια στιγμή θα έχουμε ότι

$$p_1 p_2 \cdots p_t V \cong R/\langle b'_1 \rangle \oplus \cdots \oplus R/\langle b'_r \rangle,$$

όπου $\langle b'_1 \rangle = \cdots = \langle b'_j \rangle = R$. Αν υπάρχει $\langle b'_{j+1} \rangle \neq R$ τότε η μοναδικότητα προκύπτει με επαγωγή στο πλήθος των προσθεταίων r , αφού έχουμε καταλήξει σε ένα module με μικρότερο πλήθος προσθετέων. Διαφορετικά έχουμε και πάλι το μονοσήμαντο της ανάλυσης αφού τα $b_1, \dots, b_r$ είναι όλα ίσα με $p_1 p_2 \cdots p_t$ μέχρι πολλαπλασιασμό με αντιστρέψιμο στοιχείο, συνεπώς τα ιδεώδη που παράγουν είναι ίσα. $\square$

Ρητή κανονική μορφή

Επιστρέφουμε και πάλι στην δράση του $\mathbb{F}[x]$ στο V όπως ορίστηκε στην (VI.18).

Ορισμός VI.4.18. Αν υπάρχει ένα στοιχείο $v \in V$ ώστε $V = \mathbb{F}[x]v$, τότε θα λέμε ότι το V είναι ένα κύριο $\mathbb{F}[x]$ -module.

Ένα κύριο $\mathbb{F}[x]$ -module V παράγεται ως $\mathbb{F}$ -διανυσματικός χώρος από τα στοιχεία $v, xv, x^2v, \dots$. Θεωρούμε το ελάχιστο πολυώνυμο $m_L(x)$ του L . Είναι σαφές ότι τα

$$v, xv, x^2v, \dots, x^{n-1}v \quad (\text{VI.23})$$

αποτελούν μια βάση του $V = \mathbb{F}[x]v$. Πράγματι, μια σχέση γραμμικής εξάρτησης θα έδινε ένα πολυώνυμο $g(x)$ μικρότερου βαθμού από το ελάχιστο πολυώνυμο ώστε $g(L)v = 0$. Αυτό θα έδινε ότι το $g(L) = 0$, αφού κάθε στοιχείο $w \in V$, είναι $w = h(x)v = h(L)v$ συνεπώς $g(L)w = g(L)h(L)v = h(L)g(L)v = 0$. Επίσης τα στοιχεία της εξίσωσης παράγουν τον V , αφού κάθε πολυώνυμο $f(x) \in \mathbb{F}[x]$, γράφεται ως $f(x) = q(x)m_L(x) + u(x)$, με $u(x) = 0$ ή $\deg u(x) \leq n$, άρα

$$V \ni w = f(x)v = q(x)m_L(x)v + u(x)v = u(x)v.$$

Ο πίνακας της L ως προς την παραπάνω βάση είναι σαφές ότι είναι ο

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

όπου $m_L(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{F}[x]$.

Παρατήρηση VI.4.19. Το κύριο module V είναι ισόμορφο με το $\mathbb{F}[x]/\langle m_L(x) \rangle$, όπως παρατηρεί κανείς από τον μορφισμό $\mathbb{F}[x] \rightarrow \mathbb{F}[x]v = V$ με πυρήνα $\langle m_L(x) \rangle$.

Άμεση συνέπεια του θεωρήματος VI.4.17 είναι το

Θεώρημα VI.4.20. Έστω V ένας μη-μηδενικός πεπερασμένης διάστασης διανυσματικός χώρος πάνω από το σώμα $\mathbb{F}$ και έστω $L : V \rightarrow V$ γραμμική συνάρτηση. Τότε ο V διασπάται ως ευθύ άθροισμα L -αναλλοίωτων υπόχωρων

$$V = V_1 \oplus \cdots \oplus V_r,$$

όπου κάθε V_i είναι ένα κύριο $\mathbb{F}[x] = \mathbb{F}[L]$ -module με ελάχιστο πολυώνυμο $q_i \neq 0$ ώστε $q_1 \mid q_2 \mid \cdots \mid q_r$. Η ακολουθία $\{q_1, \dots, q_r\}$ είναι μονοσήμαντα ορισμένη.

Απόδειξη. Αυτή είναι μια άμεση μεταφορά του θεωρήματος VI.4.17 στο $\mathbb{F}[x]$ -module V . □

Παρατήρηση VI.4.21. Η παραπάνω διάσπαση ονομάζεται ρητή κανονική μορφή του $L : V \rightarrow V$. Έχει το πλεονέκτημα ότι δεν απαιτεί από το σώμα $\mathbb{F}$ να είναι αλγεβρικά κλειστό.

Η κανονική μορφή Jordan μπορεί να προκύψει από την ρητή κανονική μορφή ως εξής:

Θεώρημα VI.4.22. Αν $\mu_L(x) \in \mathbb{F}[x]$ είναι της μορφής $\mu_L(x) = (x - \lambda)^k$ τότε το κύριο module $V = \mathbb{F}[x]v$ έχει μια βάση B ως $\mathbb{F}$ -διανυσματικός χώρος ώστε ο πίνακας

$$(L, B, B) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

Απόδειξη. Το V είναι ένα κύριο $\mathbb{F}[x]$ -module με ελάχιστο πολυώνυμο $(x - \lambda)^k$ και ταυτίζεται με το $\mathbb{F}[z]v$ για κάποιο $v \in V$. Τα στοιχεία

$$v, (x - \lambda)v, (x - \lambda)^2v, \dots, (x - \lambda)^{k-1}v \quad (\text{VI.24})$$

αποτελούν μια βάση του V . Είναι γραμμικά ανεξάρτητα, διότι κάθε γραμμική σχέση εξάρτησης οδηγεί σε ένα μη-μηδενικό πολυώνυμο που μηδενίζει τον L και είναι βαθμού μικρότερου του k .

Παρατηρούμε ότι αν θεωρήσουμε την διατεταγμένη βάση με την ανάποδη διάταξη από αυτή της (VI.24) δηλαδή

$$e_1 = (x - \lambda)^{k-1}v, e_2 = (x - \lambda)^{k-2}v, \dots, e_j = (x - \lambda)^{k-j}v, \dots, e_k = v$$

τότε ο πίνακας έχει ακριβώς την μορφή που θέλουμε. Πράγματι

$$xe_j = (x - \lambda)e_j + \lambda e_j = (x - \lambda)(x - \lambda)^{k-j}v + \lambda e_j = \lambda e_j + e_{j-1}.$$

Ας σημειωθεί ότι στον παραπάνω τύπο για λόγους ομοιομορφίας θέσαμε $e_0 = 0$. □

VII.1 Διανυσματικοί χώροι με νόρμα

Ορισμός VII.1.1. Ένας διανυσματικός χώρος με νόρμα είναι ένας διανυσματικός χώρος V πάνω από το σώμα $\mathbb{F}$ των πραγματικών ή μιγαδικών αριθμών, συνοδευόμενος με μια νόρμα ή απόλυτη τιμή, συμβολίζεται με $\|\cdot\| : V \rightarrow \mathbb{R}$ και ικανοποιεί τις παρακάτω ιδιότητες για κάθε διάνυσμα $u, v \in V$ και κάθε $c \in \mathbb{F}$:

1. Θετικά ορισμένη: $\|u\| \geq 0$, και $\|u\| = 0$ αν και μόνο αν $u = 0$.
2. Ομογένεια: $\|cu\| = |c|\|u\|$
3. Τριγωνική Ανισότητα: $\|u + v\| \leq \|u\| + \|v\|$

Παρατήρηση VII.1.2. Στην δεύτερη σχέση το $|c|$ είναι η απόλυτη τιμή ενός στοιχείου του $\mathbb{R}$, $c \in \mathbb{R}$ ή είναι το μέτρο ενός μιγαδικού αριθμού $|c| = \sqrt{\operatorname{Re}^2 c + \operatorname{Im}^2 c}$ αν $c \in \mathbb{C}$.

Παρατήρηση VII.1.3. Είναι σαφές από τα αξιώματα της νόρμας ότι $\|v\| \geq 0$ για κάθε $v \in V$. Πράγματι,

$$0 = \|v - v\| \leq \|v\| + \|-v\| = 2\|v\|.$$

Παρατήρηση VII.1.4. Ισχύει

$$\| \|v\| - \|u\| \| \leq \|v - u\| \text{ για κάθε } v, u \in V.$$

Πράγματι,

$$\|u\| = \|(u - v) + v\| \leq \|u - v\| + \|v\|,$$

δηλαδή $\|u\| - \|v\| \leq \|u - v\|$ και αλλάζοντας τους ρόλους των u, v έχουμε $\|v\| - \|u\| \leq \|v - u\| = \|u - v\|$ δηλαδή

$$-\|u\| + \|v\| \leq \|u - v\| \leq \|u\| - \|v\|,$$

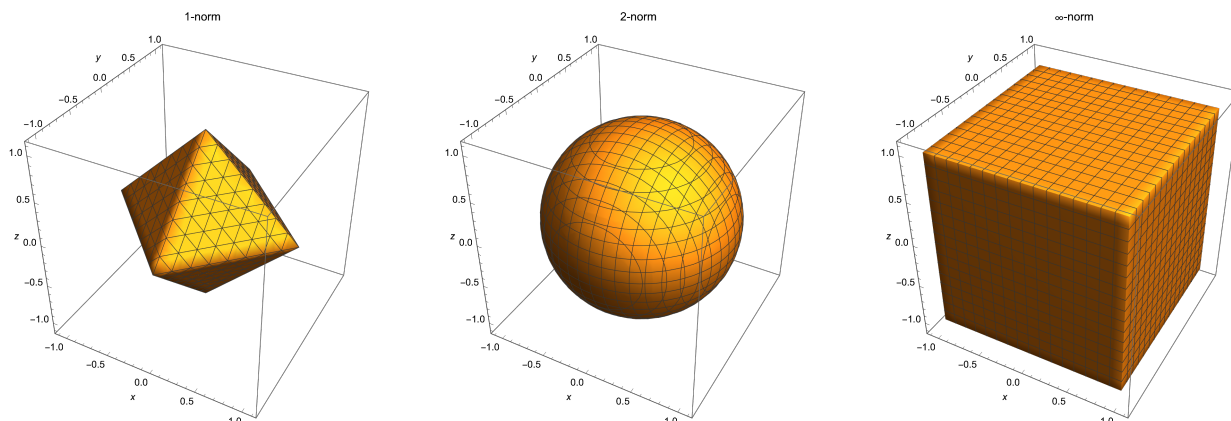
το οποίο είναι και το ζητούμενο.

Παραδείγματα VII.1.5. 1. Ο χώρος $(\mathbb{R}, \|\cdot\|)$ με $\|x\| = |x|$ για κάθε $x \in \mathbb{R}$.

2. Ο χώρος $(\mathbb{C}, \|\cdot\|)$ με $\|x\| = |x|$ για κάθε $x \in \mathbb{C}$.

3. Ο χώρος $(\mathbb{R}^n, \|\cdot\|_1)$ με

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \text{ όπου } x = (x_1, \dots, x_n).$$



Σχήμα VII.1: Οι μοναδιαίες μπάλες $B = \{x : \|x\| \leq 1\}$ στον χώρο $\mathbb{R}^3$ για $\|\cdot\| = \|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$.

4. Ο χώρος $(\mathbb{R}^n, \|\cdot\|_2)$ με

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}, \text{ όπου } x = (x_1, \dots, x_n).$$

5. Ο χώρος $(\mathbb{R}^n, \|\cdot\|_\infty)$ με

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \text{ όπου } x = (x_1, \dots, x_n).$$

6. Ο χώρος $(C[a, b], \|\cdot\|_\infty)$ των συνεχών συναρτήσεων με νόρμα

$$\|f\| = \max_{a \leq x \leq b} |f(x)|.$$

Στο παραπάνω είναι σημαντικό η f να είναι συνεχής σε κλειστό διάστημα ώστε να έχει μέγιστο.

Παρατήρηση VII.1.6. Σε ένα διανυσματικό χώρο με εσωτερικό γινόμενο μπορούμε να ορίσουμε την έννοια της σύγκλισης μιας ακολουθίας (v_n) στοιχείων $v_n \in V$ στο $v \in V$ ως εξής:

Το όριο $\lim_{n \rightarrow \infty} v_n = v$ αν και μόνο αν

Για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$, ώστε για $n > n_0$ ισχύει $\|v_n - v\| < \epsilon$.

Οι ιδιότητες του ορίου τέτοιων ακολουθιών είναι παρόμοιες με αυτές που ισχύουν για τα όρια ακολουθιών πραγματικών αριθμών γιατί ισχύουν οι ιδιότητες της νόρμας που είναι παρόμοιες με αυτές της απόλυτης τιμής, την οποία και γενικεύουν.

Έστω ένας διανυσματικός χώρος V . Επιλέγοντας μια βάση B έχουμε ένα ισομορφισμό $[\cdot]_B : V \rightarrow \mathbb{R}^n$, οπότε μπορούμε να ορίσουμε μια νόρμα ως εξής:

$$\|v\|_V = \|[v]_B\|_{\mathbb{R}^n}.$$

Δηλαδή μετράμε με μια νόρμα του $\mathbb{R}^n$ τις συντεταγμένες του διανύσματος. Η νόρμα αυτή εξαρτάται από την επιλογή της βάσης. Θα ορίσουμε μια έννοια ισοδυναμίας νορμών και θα δείξουμε στην πρόταση **VII.1.11** ότι η εξάρτηση αυτή από την βάση δεν πειράζει και πολύ αφού όλες οι νόρμες σε έναν πεπερασμένης διάστασης διανυσματικό χώρο είναι ισοδύναμες.

Ορισμός VII.1.7. Θα λέμε ότι οι νόρμες $\|\cdot\| : V \rightarrow \mathbb{R}$ και $\|\cdot\|' : V \rightarrow \mathbb{R}$ είναι ισοδύναμες αν υπάρχουν σταθερές $A > 0, B > 0$ ώστε

$$A\|v\|' \leq \|v\| \leq B\|v\|' \text{ για κάθε } v \in V.$$

Παρατήρηση VII.1.8. Η παραπάνω σχέση είναι σχέση ισοδυναμίας, δηλαδή

- Κάθε νόρμα είναι ισοδύναμη με τον εαυτό της. Πράγματι, αρκεί να διαλέξουμε $A, B = 1$.
- Αν η νόρμα $\|\cdot\|$ είναι ισοδύναμη με την $\|\cdot\|'$ δηλαδή υπάρχουν A, B ώστε να ισχύουν

$$A\|v\|' \leq \|v\| \leq B\|v\|' \Rightarrow B^{-1}\|v\| \leq \|v\|' \leq A^{-1}\|v\|$$

- Αν η $\|\cdot\|$ είναι ισοδύναμη με την $\|\cdot\|'$ είναι ισοδύναμη με την $\|\cdot\|''$ δηλαδή υπάρχουν A_1, A_2 και B_1, B_2 ώστε

$$\begin{aligned} A_1\|v\|' &\leq \|v\| \leq B_1\|v\|' \\ A_2\|v\|'' &\leq \|v\|' \leq B_2\|v\|'' \end{aligned}$$

τότε

$$A_1A_2\|v\|'' \leq A_1\|v\|' \leq \|v\| \leq B_1\|v\|' \leq B_1B_2\|v\|$$

δηλαδή και $\|\cdot\|$ είναι ισοδύναμη με την $\|\cdot\|''$.

Η έννοια της νόρμας σε ένα χώρο μας δίνει την δυνατότητα να ορίσουμε όρια.

Ορισμός VII.1.9. Θα λέμε ότι η ακολουθία (a_n) στοιχείων του V συγκλίνει στο $a \in V$ αν για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$ ώστε για κάθε $n \geq n_0$ να ισχύει $\|a_n - a\| < \epsilon$.

Θα λέμε ότι η ακολουθία (a_n) στοιχείων του V είναι Cauchy αν για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$ ώστε για κάθε $m, \ell \geq n_0$ να ισχύει $\|a_m - a_\ell\| < \epsilon$.

Τέλος ο χώρος V θα λέγεται πλήρης αν και μόνο αν κάθε ακολουθία Cauchy σε αυτόν έχει όριο στον V .

Παρατήρηση VII.1.10. Αν οι μετρικές $\|\cdot\|$ και $\|\cdot\|'$ είναι ισοδύναμες τότε αν το όριο της (a_n) ως προς την $\|\cdot\|$ υπάρχει και είναι $a \in V$ τότε το όριο υπάρχει και ως προς την $\|\cdot\|'$ και είναι επίσης a .

Πρόταση VII.1.11. Όλες οι νόρμες σε ένα πεπερασμένης διάστασης διανυσματικό χώρο πάνω από ένα πλήρες σώμα είναι ισοδύναμες.

Απόδειξη. Θεωρούμε ένα πλήρες σώμα $\mathbb{F}$ ως προς την νόρμα $|\cdot|$. Θα κάνουμε επαγωγή ως προς την διάσταση n του V . Η περίπτωση $\{0\}$ είναι τετριμμένη. Ας υποθέσουμε ότι $\dim V = 1$ και έστω $v_0 \neq 0$ ένα στοιχείο του V , οπότε $V = \mathbb{F}v_0$, οπότε $\|cv_0\| = |c|\|v_0\|$. Συνεπώς δύο νόρμες $\|\cdot\|, \|\cdot\|'$ στον μονοδιάστατο χώρο V είναι η μία πολλαπλάσιο της άλλης.

Υποθέτουμε ότι $n \geq 2$ και ότι η πρόταση ισχύει για όλους τους διανυσματικούς χώρους διάστασης $n-1$. Έστω V διανυσματικός χώρος διάστασης n . Διαλέγουμε μια $\mathbb{F}$ βάση $\{v_1, \dots, v_n\}$ του V . Για μια οποιαδήποτε νόρμα $\|\cdot\|$ του V θα δείξουμε ότι είναι ισοδύναμη με την νόρμα $\|\cdot\|_\infty$ η οποία ορίζεται ως

$$\left\| \sum_{i=1}^n c_i v_i \right\| = \max_{1 \leq i \leq n} |c_i|.$$

Παρατηρούμε ότι ο V ως προς την $\|\cdot\|_\infty$ είναι πλήρης διότι αν έχουμε μια ακολουθία Cauchy (w_n) στον V ως προς την $\|\cdot\|_\infty$ αυτό σημαίνει ότι για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$, ώστε για $\ell, m > n_0$ έχουμε $\|a_\ell - a_m\| < \epsilon$. Γράφοντας τα a_m ως προς την βάση $v_1, \dots, v_n$ έχουμε ότι

$$a_m = c_{1m}v_1 + \dots + c_{nm}v_n$$

οπότε ο ορισμός της $\|\cdot\|_\infty$ δίνει ότι οι ακολουθίες $(c_{im})_{m \in \mathbb{N}}$ είναι Cauchy ως προς την $|\cdot|$, και λόγω πληρότητας υπάρχει το όριο $c_i = \lim_{m \rightarrow \infty} c_{im}$. Τότε το $a = \sum c_i v_i$ είναι το όριο της (a_m) στον V :

$$\|a - a_m\|_\infty = \max_{1 \leq i \leq n} |c_i - c_{im}| \rightarrow 0.$$

Θα πρέπει να βρούμε σταθερές ώστε

$$A\|v\|_\infty \leq \|v\| \leq B\|v\|_\infty.$$

Το τυχαίο $v \in V$ το γράφουμε $v = c_1v_1 + \dots + c_nv_n$. Έχουμε

$$\|v\| \leq \sum_{i=1}^n \|c_i v_i\| = \sum_{i=1}^n |c_i| \|v_i\| \leq B \max |c_i| = B\|v\|_\infty,$$

όπου $B = \sum \|v_i\| > 0$.

Θα δείξουμε τώρα ότι υπάρχει $A > 0$ ώστε $A\|v\|_\infty \leq \|v\|$. Ας υποθέσουμε ότι αυτό δεν είναι σωστό για κανένα A , οπότε για κάθε $A = 1/k$, $k = 1, 2, \dots$ υπάρχει ένα $u_k \in V$, ώστε

$$\|u_k\| < (1/k)\|u_k\|_\infty.$$

Από τον ορισμό του $\|\cdot\|_\infty$ σε όρους της βάσης $v_1, \dots, v_n$ έχουμε ότι το $\|u_k\|_\infty$ είναι η απόλυτη τιμή κάποιας συντεταγμένης του u_k . Αφού το σύνολο u_k είναι άπειρο και υπάρχουν πεπερασμένα το πλήθος στοιχεία της βάσης $v_1, \dots, v_n$, υπάρχει κάποιος δείκτης i , $1 \leq i \leq n$ για τον οποίο ισχύει ότι άπειρα u_k έχουν $\|u_k\|$ ίσο με την συντεταγμένη που αντιστοιχεί στο v_i στοιχείο της βάσης. Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι $i = n$. Παρατηρούμε ότι αν πολλαπλασιάσουμε το u_k με κάποιο μη μηδενικό στοιχείο του $\mathbb{F}$ η ανισότητα $\|u_k\| < (1/k)\|u_k\|_\infty$ παραμένει σε ισχύ, αφού και τα δύο μέλη της πολλαπλασιάζονται με το ίδιο στοιχείο.

Έτσι μπορούμε, αφού περάσουμε σε κάποια υποακολουθία $\{u_{k_j}\}$ να υποθέσουμε ότι

1. $\|u'_{k_j}\|_\infty = 1$,
2. Η συντεταγμένη του v_n του u'_{k_j} να είναι ίση με 1.
3. $\|u'_{k_j}\| < (1/k_j)\|v'_n\|_\infty = 1/k_j$.

Αφού το $k_j \rightarrow \infty$ όταν $j \rightarrow \infty$ έχουμε ότι $\|u'_{k_j}\| \rightarrow 0$, όταν $j \rightarrow \infty$. Θέτουμε $w_j = u'_{k_j} - v_n$, οπότε w_j ανήκει στον χώρο $W = \langle v_1, \dots, v_n \rangle$ και έχουμε

$$\|w_j + v_n\| = \|u'_{k_j}\| \rightarrow 0 \quad (\text{VII.1})$$

Θεωρούμε τις διαφορές

$$\|w_j - w_{j'}\| = \|(w_j + v_n) - (w_{j'} + v_n)\| \leq \|(w_j + v_n)\| - \|(w_{j'} + v_n)\| \rightarrow 0. \quad (\text{VII.2})$$

Συνεπώς η ακολουθία w_j είναι ακολουθία Cauchy στον W ως προς την $\|\cdot\|$ περιορισμένη στον W .

Αφού ο χώρος W έχει διάσταση $n - 1$ από την επαγωγική υπόθεση όλες οι νόρμες είναι ισοδύναμες σε αυτόν, συνεπώς $\|\cdot\|_\infty \leq C\|\cdot\|$ και $\|\cdot\| \leq C'\|\cdot\|_\infty$ στον W για κάποια $C > 0$, $C' > 0$. Από την (VII.2) έχουμε ότι

$$\|w_j - w_{j'}\|_\infty \leq C\|w_j - w_{j'}\| \rightarrow 0.$$

Συνεπώς η ακολουθία w_j είναι ακολουθία Cauchy στον W ως προς την $\|\cdot\|_\infty$ περιορισμένη στον W . Ο χώρος $(W, \|\cdot\|_\infty)$ είναι πλήρης οπότε υπάρχει ένα όριο w για τα w_j ως προς την $\|\cdot\|_\infty$, δηλαδή

$$\|w - w_j\|_\infty \rightarrow 0.$$

Τότε $\|w - w_j\| \leq C'\|w - w_j\|_\infty \rightarrow 0$ συνεπώς $\|w - w_j\| \rightarrow 0$. Σε συνδυασμό με την (VII.1) έχουμε ότι

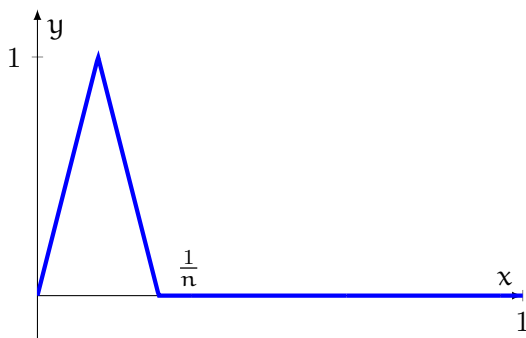
$$\|w + v_n\| = \|(w - w_j) + (w_j + v_n)\| \leq \|w - w_j\| + \|w_j + v_n\| \rightarrow 0$$

οπότε $w = -v_n$, όμως $w \in W$ και $-v_n \notin W$, από όπου έχουμε μια αντίφαση. $\square$

Παρατήρηση VII.1.12. Η πρόταση VII.1.11 δεν ισχύει αν η διάσταση του V δεν είναι πεπερασμένη. Για παράδειγμα στον χώρο $C([0, 1])$ των συνεχών συναρτήσεων $f: [0, 1] \rightarrow \mathbb{R}$ ορίζονται οι νόρμες

$$\|f\|_\infty = \max_{0 \leq x \leq 1} |f(x)| \quad \|f\|_1 = \int_0^1 |f(x)| dx$$

Είναι σαφές ότι $\|f\|_1 \leq \|f\|_\infty$ αλλά δεν υπάρχει C ώστε $\|f\| \leq C\|f\|_1$. Μπορούμε να διαλέξουμε τις συναρτήσεις $f_n(x)$ οι οποίες να είναι 0 στο διάστημα $(1/n, 1)$ και να έχουν το τρίγωνο σχήμα με κορυφή 1 στο $1/2n$ όπως στο παρακάτω σχήμα:



Η ακολουθία f_n συγκλίνει στην μηδενική συνάρτηση με την $\|\cdot\|_1$ νόρμα αφού το εμβαδόν της διαφοράς τείνει στο 0 αλλά όχι με την $\|\cdot\|_\infty$ αφού η νόρμα $\|\cdot\|_\infty$ βλέπει την μεγαλύτερη τιμή της $f_n(x)$ που είναι 1.

VII.2 Παραγωγή συναρτήσεων

Στην παράγραφο I.8.7 είδαμε ένα ορισμό της παραγώγου συνάρτησης $f: \mathbb{R} \rightarrow \mathbb{R}$. Τον ορισμό αυτό θέλουμε να τον επεκτείνουμε σε ένα ορισμό παραγώγου για μια συνάρτηση $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Ακολουθώντας την παρουσίαση στο [1] και μιμούμενοι στην ουσία την ιδέα για τον ορισμό για την μία μεταβλητή ορίζουμε

Ορισμός VII.2.1. Έστω $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ και $a \in \mathbb{R}^n$. Θα λέμε ότι η f είναι διαφορίσιμη στο a αν υπάρχει μια συνάρτηση $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$, η οποία να είναι συνεχής στο a και να ικανοποιεί

$$f(x) - f(a) = \phi(x)(x - a). \quad (\text{VII.3})$$

Το $\phi(a) \in \mathbb{R}^{m,n}$ θα τον ονομάζουμε παράγωγο της f στο a , και θα το συμβολίζουμε με $Df(a)$. Την συνάρτηση ϕ θα την ονομάζουμε συνάρτηση κλίσης.

Θα χρειαστεί να εισάγουμε νόρμες οι οποίες σέβονται τον πολλαπλασιασμό πινάκων.

Ορισμός VII.2.2. Μια νόρμα στον χώρο των $n \times n$ πινάκων θα λέγεται υποπολλαπλασιαστική αν και μόνο αν ισχύει

$$\|A \cdot B\| \leq \|A\| \|B\|.$$

Επίσης, θα λέμε ότι η νόρμα πινάκων είναι συμβατή με την νόρμα του διανυσματικού χώρου αν ισχύει

$$\|Av\| \leq \|A\| \|v\|.$$

Ορισμός VII.2.3. Έστω $\mathbb{F} = \mathbb{C}$ ή $\mathbb{F} = \mathbb{R}$. Από κάθε νόρμα διανυσματικών χώρων $\|\cdot\| : \mathbb{F}^n \rightarrow \mathbb{R}$ μπορούμε να ορίσουμε την *επαγόμενη νόρμα*

$$\|A\| = \max_{v \in \mathbb{F}^n \setminus \{0\}} \frac{\|Av\|}{\|v\|} = \max_{v \in \mathbb{F}^n : \|v\|=1} \|Av\|. \quad (\text{VII.4})$$

Λήμμα VII.2.4. Η εξίσωση (VII.4) ορίζει νόρμα συμβατή και υποπολλαπλασιαστική. Επιπλέον $\|I_n\| = 1$.

Απόδειξη. Από τον ορισμό της νόρμας έχουμε για κάθε $v \in \mathbb{R}^n$, $v \neq 0$

$$\|Av\|/\|v\| \leq \|A\|$$

από όπου προκύπτει ότι η νόρμα είναι συμβατή. Η περίπτωση $v = 0$ είναι προφανής.

Παρατηρούμε, χρησιμοποιώντας την συμβατότητα της νόρμας, ότι για κάθε $v \in \mathbb{R}^n$

$$\|ABv\| \leq \|A\| \|Bv\| \leq \|A\| \|B\| \|v\|$$

από όπου έχουμε ότι

$$\|AB\| \leq \|A\| \|B\|$$

από όπου έχουμε ότι η νόρμα είναι υποπολλαπλασιαστική.

Τέλος η νόρμα του $\|I_n\| = 1$ από τον ορισμό αφού το μέγιστο υπολογίζεται πάνω στην σταθερή ποσότητα $\|v\|/\|v\| = 1$. $\square$

Ασκήσεις από το 2020_03_25.pdf

Παρατήρηση VII.2.5. Επιστρέφοντας στον ορισμό της παραγώγου παρατηρούμε ότι αν υπάρχει μία συνάρτηση κλίσης ϕ τότε

$$\frac{\|f(x) - f(a) - \phi(a)(x - a)\|}{\|x - a\|} = \frac{\|(\phi(x) - \phi(a))\|}{\|x - a\|} \leq \|\phi(x) - \phi(a)\|$$

και η συνέχεια της ϕ μας δίνει ότι

$$\lim_{x \rightarrow a} \frac{\|f(x) - f(a) - \phi(a)(x - a)\|}{\|x - a\|} = 0.$$

Παρατήρηση VII.2.6. Η συνεχής συνάρτηση ϕ της οποίας την ύπαρξη απαιτήσαμε παραπάνω μπορεί να μην είναι μοναδική. Για παράδειγμα, αν $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ορίζεται από $f(x, y) = xy$, και $a = (x_0, y_0) \in \mathbb{R}^2$ τότε

$$\begin{aligned} f(x, y) - f(x_0, y_0) &= (y_0, x)(x - x_0, y - y_0)^t \\ &= (y, x_0)(x - x_0, y - y_0)^t \end{aligned}$$

, δηλαδή και η $\phi_1(x, y) = (y_0, x)$ και η $\phi_2(x, y) = (y, x_0)$ ικανοποιούν την απαίτηση (VII.3). Παρατηρούμε ότι $\phi_1(x_0, y_0) = \phi_2(x_0, y_0)$, οπότε κάνοντας χρήση είτε της μίας είτε της άλλης παίρνουμε την ίδια απάντηση για το $Df(a)$.

Πρόταση VII.2.7. Η παράγωγος $Df(a)$ είναι καλά ορισμένη, δηλαδή αν $\phi_i : \mathbb{R}^n \rightarrow \mathbb{R}^{m,n}$, $i = 1, 2$ είναι δύο συναρτήσεις κλίσης τότε $\phi_1(a) = \phi_2(a)$.

Απόδειξη. Για δύο συναρτήσεις κλίσης ϕ_1, ϕ_2 θεωρούμε την διαφορά τους $\eta(x) = \phi_1(x) - \phi_2(x)$. Από τον ορισμό των συναρτήσεων κλίσης έχουμε ότι

$$\eta(x)(x - a) = 0,$$

συνεπώς

$$\|\eta(a)(x - a)\| = \|(\eta(a) - \eta(x))(x - a)\| \leq \|\eta(a) - \eta(x)\| \cdot \|x - a\|$$

το οποίο οδηγεί στο

$$\|\eta(x)\| = \left\| \eta(a) \left(\frac{x - a}{\|x - a\|} \right) \right\| \leq \|\eta(a) - \eta(x)\|.$$

Η συνέχεια της η στο a , εξασφαλίζει ότι $\eta(a) = 0$, δηλαδή το ζητούμενο. $\square$

Θα δείξουμε τώρα μερικές ιδιότητες της παραγώγου με βάση τον ορισμό που δώσαμε:

Θεώρημα VII.2.8. Αν $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ είναι διαφορίσιμες, τότε το ίδιο είναι και $\lambda f + \mu g$ και ισχύει ότι $D(\lambda f + \mu g)(a) = \lambda Df(a) + \mu Dg(a)$.

Απόδειξη. Αν ϕ, ψ είναι συναρτήσεις κλίσης για τις συναρτήσεις f, g αντίστοιχα τότε

$$(\lambda f + \mu g)(x) - (\lambda f + \mu g)(a) = \lambda(f(x) - f(a)) + \mu(g(x) - g(a)) = (\lambda\phi + \mu\psi)(x)(x - a).$$

Όμως $\lambda\phi + \mu\psi$ είναι συνεχής και παίζει τον ρόλο της συνάρτησης κλίσης για την $\lambda f + \mu g$. Αυτό αποδεικνύει και ότι

$$D(\lambda f + \mu g)(a) = \lambda\phi(a) + \mu\psi(a) = \lambda Df(a) + \mu Dg(a).$$

$\square$

Θεώρημα VII.2.9 (Κανόνας Αλυσίδας). Αν η $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ είναι διαφορίσιμη στο a και η $g : \mathbb{R}^m \rightarrow \mathbb{R}^l$ είναι διαφορίσιμη στο $f(a)$, τότε η $g \circ f$ είναι διαφορίσιμη στο a και ισχύει:

$$D(g \circ f) = Dg(f(a))Df(a).$$

Απόδειξη. Έχουμε το παρακάτω διάγραμμα συναρτήσεων:

$$\begin{array}{ccccc} & & g \circ f & & \\ & \nearrow & & \searrow & \\ \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^m & \xrightarrow{g} & \mathbb{R}^l \end{array}$$

Παρατηρούμε ότι $Dg(f(a)) \in \mathbb{R}^{l,m}$ και $Df(a) \in \mathbb{R}^{m,n}$, συνεπώς το $Dg(f(a))Df(a)$ είναι ένας καλά ορισμένος πολλαπλασιασμός πινάκων.

Έστω $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^{m,n}$ και $\psi : \mathbb{R}^m \rightarrow \mathbb{R}^{l,m}$ συναρτήσεις κλίσης για τις f, g . Υπολογίζουμε ότι

$$g(f(x)) - g(f(a)) = \psi(f(x))(f(x) - f(a)) = \psi(f(x))\phi(x)(x - a).$$

Παρατηρούμε ότι η $\psi(f(x))\phi(x)$ είναι συνεχής, ως σύνθεση και στην συνέχεια γινόμενο συνεχών συναρτήσεων και συνεπώς παίζει τον ρόλο της συνάρτησης κλίσης. Για το τελικό αποτέλεσμα υπολογίζουμε

$$D(\psi(f(x))\phi(x)) = \psi(f(a))\phi(a) = Dg(f(a))Df(a).$$

$\square$

VII.2.1 Συναρτήσεις $\mathbb{R}^n \rightarrow \mathbb{R}$

Θεωρούμε μια παραγωγίσιμη στο x_0 συνάρτηση $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Υπάρχει μία συνεχής συνάρτηση κλίσης $(g_1^{(x_0)}(x), \dots, g_n^{(x_0)}(x))$ ώστε

$$f(x) = f(x_0) = (g_1^{(x_0)}(x), \dots, g_n^{(x_0)}(x))(x - x_0),$$

όπου $x = (x_1, \dots, x_n)^t$ και $x_0 = (x_1^{(0)}, \dots, x_n^{(0)})$. Έχουμε τοποθετήσει το (x_0) στον εκθέτη για να θυμόμαστε ότι η συνάρτηση κλίσης εξαρτάται από το x_0 .

Οι τιμές $g_i^{(x_0)}(x_0)$ είναι οι μερικές παράγωγοι της f υπολογισμένες στο x_0 και γράφονται ως $\frac{\partial f}{\partial x_i}$.

Πρόταση VII.2.10. Θεωρούμε την συνάρτηση $f: \mathbb{R}^n \subset U \rightarrow \mathbb{R}$ η οποία ορίζεται στο καρτεσιανό γινόμενο ανοιχτών διαστημάτων $U = (a_1, b_1) \times (a_2, b_2) \times \dots \times (a_n, b_n)$. Αν η f είναι διαφορίσιμη στο $x_0 \in U$ και $f(x_0)$ είναι μέγιστο ή ελάχιστο της f στο U τότε $Df(x_0) = 0$.

Απόδειξη. Υποθέτουμε ότι $f(x_0) \leq f(x)$ για κάθε $x \in U$ και ότι η f διαφορίσιμη στο x_0 . Τότε

$$0 \leq f(x) - f(x_0) = \phi(x)(x - x_0), \quad (\text{VII.5})$$

για όλα τα $x \in U$ όπου η ϕ είναι μια συνάρτηση κλίσης της f στο x_0 . Σταθεροποιούμε το $h \in \mathbb{R}^n$ και θεωρούμε ένα αρκετά μικρό ϵ ώστε $x_0 + th \in U$ για κάθε $t \in (-\epsilon, \epsilon)$. Από την (VII.5) έχουμε ότι

$$t\phi(x_0 + th)h \geq 0$$

για κάθε $t \in (-\epsilon, \epsilon)$. Συνεπώς,

$$\phi(x_0 + th)h \geq 0 \text{ αν } t > 0,$$

και

$$\phi(x_0 + th)h \leq 0 \text{ αν } t < 0.$$

Η συνέχεια της ϕ μας εξασφαλίζει ότι $\phi(x_0)h = 0$ και επειδή το h είναι τυχαίο έχουμε ότι $D(f_{x_0}) = 0$. $\square$

Υποθέτουμε ότι κάθε συνάρτηση $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ είναι και αυτή παραγωγίσιμη στο x_0 άρα γράφεται ως

$$g_i(x) = g_i(x_0) + (h_{i,1}^{(x_0)}(x), \dots, h_{i,n}^{(x_0)}(x))(x - x_0),$$

όπου οι $h_{i,j}$ είναι όλες παραγωγίσιμες συναρτήσεις.

Μπορούμε λοιπόν να γράψουμε

$$f(x) = f(x_0) + (g_1^{(x_0)}(x_0), \dots, g_n^{(x_0)}(x_0))(x - x_0) + \sum_{i,j=1}^n h_{i,j}^{(x_0)}(x) (x_i - x_i^{(0)}) (x_j - x_j^{(0)}).$$

Ορισμός VII.2.11. Ορίζουμε $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = h_{ij}^{(x_0)}(x_0)$. Ο $n \times n$ πίνακας $H(x_0) = (h_{i,j}^{(x_0)}(x_0))$ ονομάζεται Hessian της f .

Πρόταση VII.2.12. Ο πίνακας H είναι συμμετρικός, δηλαδή $h_{i,j}^{(x_0)}(x_0) = h_{j,i}^{(x_0)}(x_0)$.

Απόδειξη. Για μια συνάρτηση $g: \mathbb{R}^n \rightarrow \mathbb{R}$, θα συμβολίζουμε με

$$\Delta_{i,h}g(x_0) = \frac{g(x_0 + he_i) - g(x_0)}{h},$$

όπου $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots$, $e_n = (0, \dots, 0, 1)$. Ισχύει ότι

$$\Delta_{i,h}f(x_0) = g_i^{(x_0)}(x_0 + he_i), \text{ και } \lim_{h \rightarrow 0} \Delta_{i,h}f(x_0) = \frac{\partial f}{\partial x_i}(x_0).$$

Εφαρμόζουμε άλλη μια φορά τον $\Delta_{j,h'}$

$$\Delta_{j,h'}\Delta_{i,h}f(x_0) = h_{i,j}^{(x_0+he_i)}(x_0 + he_i + h'e_j) \xrightarrow{h' \rightarrow 0} \frac{\partial^2}{\partial x_i \partial x_j} f(x_0 + he_i).$$

Ομοίως έχουμε ότι

$$\Delta_{i,h}\Delta_{j,h'}f(x_0) = h_{j,i}^{(x_0+h'e_j)}(x_0 + he_i + h'e_j) \xrightarrow{h \rightarrow 0} \frac{\partial^2}{\partial x_j \partial x_i} f(x_0 + h'e_j).$$

□

Επίσης παρατηρούμε ότι

$$\Delta_{j,h'}\Delta_{i,h}f(x_0)f = \frac{1}{hh'} (f(x_0 + e_i h + e_j h') - f(x_0 + e_i h) - f(x_0 - e_j h') + f(x_0)) = \Delta_{i,h}\Delta_{j,h'}f(x_0)f.$$

Συνεπώς αφού οι δεύτερες μερικές παράγωγοι είναι συνεχείς το ζητούμενο προκύπτει ως

$$\begin{aligned} \lim_{h,h' \rightarrow 0} \Delta_{j,h'}\Delta_{i,h}f &= \lim_{h,h' \rightarrow 0} \Delta_{j,h}\Delta_{i,h'}f \Rightarrow \lim_{h \rightarrow 0} \frac{\partial^2}{\partial x_i \partial x_j} f(x_0 + he_i) = \lim_{h' \rightarrow 0} \frac{\partial^2}{\partial x_j \partial x_i} f(x_0 + h'e_i) \\ &\Rightarrow \frac{\partial^2}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2}{\partial x_j \partial x_i}(x_0). \end{aligned}$$

Συνοψίζοντας μπορούμε να εκφράσουμε μια συνάρτηση $f: \mathbb{R}^n \rightarrow \mathbb{R}$ που είναι παραγωγίσιμη με συνεχείς δεύτερες μερικές παραγώγους ως εξής:

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + (x - x_0)^t H(x_0)(x - x_0) + \text{όροι μεγαλύτερης τάξης}, \quad (\text{VII.6})$$

όπου ο πίνακας $H(x_0)$ είναι συμμετρικός. Θα δούμε στην παράγραφο **VIII.1.3** ένα κριτήριο για το είδος της συνάρτησης σε ένα κρίσιμο σημείο $Df(x_0) = 0$.

VII.3 Χώροι με εσωτερικό γινόμενο

Στο κεφάλαιο αυτό θα περιορίσουμε το σώμα $\mathbb{F}$ στην περίπτωση που $\mathbb{F} = \mathbb{R}$ ή $\mathbb{F} = \mathbb{C}$.

Ορισμός VII.3.1. Ένας χώρος με εσωτερικό γινόμενο είναι ένας διανυσματικός χώρος V πάνω από το σώμα $\mathbb{F}$, με ένα εσωτερικό γινόμενο που συμβολίζεται ως

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$$

για κάθε όλα τα $u, v, w \in V$ και κάθε $c \in \mathbb{F}$

1. Γραμμικότητα ως προς το πρώτο όρισμα: $\langle cu, v \rangle = c \langle u, v \rangle$
2. Συζυγής Συμμετρία: $\langle u, v \rangle = \overline{\langle v, u \rangle}$
3. Γραμμικότητα ως προς το δεύτερο όρισμα: $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$

4. Θετικά ορισμένη: $\langle u, u \rangle \geq 0$, και $\langle u, u \rangle = 0$ αν και μόνο αν $u = 0_V$.

Παραδείγματα VII.3.2. 1. Θεωρούμε τα στοιχεία $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Ορίζουμε την συνάρτηση

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle = \sum_{i=1}^n x_i y_i, \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στον $\mathbb{R}^n$.

2. Θεωρούμε τα στοιχεία $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{C}^n$. Ορίζουμε την συνάρτηση

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (x, y) &\longmapsto \langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i, \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στον $\mathbb{R}^n$. Αυτό στην βιβλιογραφία είναι γνωστό και ως ερμητιανό εσωτερικό γινόμενο.

Γιατί δεν ορίσαμε και στους μιγαδικούς αριθμούς το εσωτερικό γινόμενο όπως και στους πραγματικούς; Δηλαδή

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

και βάλαμε μια συζυγία στα y_i ; Από τον ορισμό του εσωτερικού γινομένου χρειαζόμαστε

$$\langle x, x \rangle = \sum_{i=1}^n x_i^2 \geq 0$$

το οποίο δεν μπορεί να εξασφαλιστεί αν $x_i \in \mathbb{C}$ αφού τα τετράγωνα μπορεί να είναι αρνητικά. Ούτε μπορούμε να εξασφαλίσουμε με αυτόν τον ορισμό ότι

$$\langle x, x \rangle = 0 \Leftrightarrow x = 0.$$

Αντιθέτως για το ερμητιανό εσωτερικό γινόμενο έχουμε

$$\langle x, x \rangle = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n |x_i|^2$$

από το οποίο προκύπτουν οι ζητούμενες ιδιότητες του γινομένου.

3. Για τους πραγματικούς αριθμούς $a < b$ θεωρούμε τον χώρο $C([a, b])$ των συνεχών συναρτήσεων $f : [a, b] \rightarrow \mathbb{R}$. Ορίζουμε την συνάρτηση

$$\begin{aligned} C([a, b]) \times C([a, b]) &\longrightarrow \mathbb{R} \\ (f, g) &\longmapsto \int_a^b f(t)g(t)dt, \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στο $C([a, b])$.

4. Θεωρούμε τον χώρο $X := \{f : [a, b] \rightarrow \mathbb{C}, f \text{ συνεχής}\}$. Ορίζουμε την συνάρτηση

$$\begin{aligned} X \times X &\longrightarrow \mathbb{C} \\ (f, g) &\longmapsto \int_a^b f(t)\overline{g(t)}dt \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στο X .

Ο **Charles Hermite** (24 Δεκεμβρίου 1822 - 14 Ιανουαρίου 1901) ήταν Γάλλος Μαθηματικός που εργάστηκε στην Θεωρία Αριθμών, τις τετραγωνικές μορφές, την θεωρία των Ελλειπτικών συναρτήσεων και γενικότερα στην Άλγεβρα. Μια αναπηρία στο δεξί του πόδι τον εμπόδισε να συνεχίσει τις σπουδές του στην περίφημη École Polytechnique. Ανέπτυξε έντονη ερευνητική δραστηριότητα και το 1856 έγινε μέλος της Γαλλικής Ακαδημίας επιστημών. Απλοποίησε την απόδειξη του Abel σχετικά με την αδύναμια επίλυσης της πεπτοβάθμιας εξίσωσης με ριζικά και κατάφερε να επιλύσει την πεπτοβάθμια εξίσωση, όχι με ριζικά αλλά με την βοήθεια ελλειπτικών συναρτήσεων. Το 1862 εντάχθηκε ξανά στην École Polytechnique ως μέλος του διδακτικού προσωπικού ενώ του 1869 έγινε καθηγητής. Είναι επίσης γνωστός για την απόδειξη του ότι ο αριθμός e είναι υπερβατικός, ενώ η μέθοδός του ακολουθήθηκε από τον Lindemann στην απόδειξη του ότι και ο π είναι υπερβατικός.



Ασκήσεις

Άσκηση VII.1 Έστω $V = \mathbb{C}^{n,n}$. Η συνάρτηση $V \times V \rightarrow V$, $(A, B) \mapsto \langle A, B \rangle = \text{tr}(AB^*)$, όπου $B^* = \overline{B}^t$ είναι ένα εσωτερικό γινόμενο. Αν $B = (b_{ij})$ τότε $\overline{B}$ είναι ο πίνακας $(\overline{b}_{ij})$, δηλαδή έχουμε εφαρμόσει την μιγαδική συζυγία σε κάθε στοιχείο του πίνακα B .

Άσκηση VII.2 Θεωρούμε το σύνολο $V = \mathbb{C}^n$ το οποίο το ταυτίζουμε με πίνακες στήλες. Για κάθε αντιστρέψιμο πίνακα $A \in \mathbb{C}^{n,n}$ θεωρούμε την συνάρτηση $V \times V \rightarrow V$, $(x, y) \mapsto \langle x, y \rangle_A = x^* A^* A y$. Δείξτε ότι είναι ένα εσωτερικό γινόμενο.

Άσκηση VII.3 Δίνεται ένα εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ σε ένα διανυσματικό χώρο με βάση $B = \{v_1, \dots, v_n\}$. Να αποδειχθεί ότι οι τιμές $\langle v_i, v_j \rangle$ καθορίζουν πλήρως το εσωτερικό γινόμενο. Επιπλέον δείξτε ότι αν $[x]_B = (x_1, \dots, x_n)^t$, $[y]_B = (y_1, \dots, y_n)^t$ τότε

$$\langle x, y \rangle = \overline{y}^t Q x = y^* Q x, \quad (\text{VII.7})$$

όπου ο Q είναι ο πίνακας $Q = (q_{ij})$ με $q_{ij} = \langle v_j, v_i \rangle$.

Άσκηση VII.4 Δείξτε ότι ικανή και αναγκαία συνθήκη για να ορίζεται ένα εσωτερικό γινόμενο από μια σχέση στις συντεταγμένες όπως δίνεται στην εξίσωση (VII.7) είναι να ισχύει $Q = Q^*$ και $x^* Q x > 0$ όταν $x \neq 0$.

Άσκηση VII.5 Αποδείξτε ότι ο τύπος

$$\left\langle \sum_{i=0}^n a_i x^i, \sum_{j=0}^m b_j x^j \right\rangle = \sum_{i=0}^n \sum_{j=0}^m \frac{a_i b_j}{i+j+1},$$

ορίζει ένα εσωτερικό γινόμενο στον χώρο των πολυωνύμων $\mathbb{R}[x]$.

VII.4 Ανισότητα Cauchy-Schwartz

Ορισμός VII.4.1. Σε ένα χώρο με εσωτερικό γινόμενο $(X, \langle \cdot, \cdot \rangle)$ ορίζεται μια συνάρτηση

$$\begin{aligned}\| \cdot \| : X &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle x, x \rangle^{1/2}\end{aligned}$$

Πρόταση VII.4.2 (Ανισότητα Cauchy-Schwarz). Σε ένα χώρο με εσωτερικό γινόμενο $(X, \langle \cdot, \cdot \rangle)$ ισχύει η ανισότητα των Cauchy-Schwarz

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|, \text{ για κάθε } x, y \in X. \quad (\text{VII.8})$$

Απόδειξη. Αν $y = 0$ τότε η ανισότητα (VII.8) ισχύει προφανώς. Υποθέτουμε ότι $y \neq 0$. Για ένα $\lambda \in \mathbb{F}$ ισχύει $\langle x - \lambda y, x - \lambda y \rangle \geq 0$. Υπολογίζουμε

$$\begin{aligned}\langle x - \lambda y, x - \lambda y \rangle &= \langle x, x - \lambda y \rangle - \lambda \langle y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda \bar{\lambda} \langle y, y \rangle \\ &= \langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \overline{\langle x, y \rangle} + \lambda \bar{\lambda} \langle y, y \rangle,\end{aligned}$$

συνεπώς $\langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \overline{\langle x, y \rangle} + \lambda \bar{\lambda} \langle y, y \rangle \geq 0$. Επιλέγουμε $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$ οπότε καταλήγουμε στο

$$\langle x, x \rangle - \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle} \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \overline{\langle x, y \rangle} + \frac{\langle x, y \rangle}{\langle y, y \rangle} \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle} \langle y, y \rangle$$

και συνεπώς

$$\langle x, x \rangle - \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle} \langle x, y \rangle \geq 0 \Rightarrow \langle x, y \rangle \overline{\langle x, y \rangle} \leq \langle x, x \rangle \langle y, y \rangle$$

ή ισοδύναμα

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \langle y, y \rangle^{1/2} = \|x\| \cdot \|y\|.$$

□

Παρατήρηση VII.4.3. Αν είχαμε ένα πραγματικό εσωτερικό γινόμενο θα μπορούσαμε απλούστερα να γράψουμε την σχέση

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle.$$

Η παραπάνω σχέση είναι θετική για κάθε $\lambda \in \mathbb{R}$ οπότε η τετραγωνική εξίσωση στο λ έχει αρνητική διακρίνουσα δηλαδή

$$\langle x, y \rangle^2 - \langle x, x \rangle \langle y, y \rangle \leq 0$$

και η ανισότητα (VII.8) έπεται.

Παρατήρηση VII.4.4. Οι ανισότητες

$$\left| \sum_{i=1}^n x_i \bar{y}_i \right| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^n |y_i|^2 \right)^{1/2}$$

για μιγαδικούς αριθμούς $x_i, y_i, i = 1, 2, \dots, n$

$$\left| \int_a^b f(t)g(t)dt \right| \leq \left(\int_a^b |f(t)|^2 dt \right)^{1/2} \left(\int_a^b |g(t)|^2 dt \right)^{1/2}$$

για συνεχείς πραγματικές συναρτήσεις είναι ειδικές περιπτώσεις της ανισότητας Cauchy-Schwarz στους χώρους $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ και $(C([a, b]), \langle \cdot, \cdot \rangle)$.

Πρόταση VII.4.5. Για ένα χώρο $(X, \langle \cdot, \cdot \rangle)$ με εσωτερικό γινόμενο η συνάρτηση $\| \cdot \| : X \rightarrow \mathbb{R}$ που ορίστηκε στον ορισμό VII.4.1 είναι μια νόρμα.

Απόδειξη. Είναι σαφές ότι $\|x\| = 0$ αν και μόνο αν $\langle x, x \rangle = 0$ αν και μόνο αν $x = 0$. Επίσης είναι σαφές ότι

$$\|\lambda x\| = \langle \lambda x, \lambda x \rangle^{1/2} = (\lambda \bar{\lambda})^{1/2} \langle x, x \rangle^{1/2} = |\lambda| \|x\|.$$

Για να αποδείξουμε την τριγωνική ανισότητα έχουμε

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \\ &= \|x\|^2 + 2\operatorname{Re}\langle x, y \rangle + \|y\|^2 \\ &\stackrel{\text{(VII.8)}}{\leq} \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \end{aligned}$$

από την οποία έπεται η τριγωνική ανισότητα. □

Σε ένα χώρο με εσωτερικό γινόμενο ισχύει το Πυθαγόρειο θεώρημα

Θεώρημα VII.4.6.

$$\langle x, y \rangle = 0 \Rightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2 \text{ για κάθε } x, y \in X \quad (\text{VII.9})$$

ή γενικότερα

$$\text{Αν για κάθε } i \neq j \text{ έχουμε } \langle x_i, x_j \rangle = 0 \text{ τότε } \|x_1 + x_2 + \dots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_n\|^2. \quad (\text{VII.10})$$

Απόδειξη. Έχουμε

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2$$

η οποία για $\langle x, y \rangle = 0$ δίνει την εξίσωση (VII.9). Θα αποδείξουμε την γενίκευση της με επαγωγή, την έχουμε ήδη αποδείξει στην περίπτωση $n = 2$. Θεωρούμε ότι ισχύει στην περίπτωση n η

$$\|x_1 + x_2 + \dots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_n\|^2$$

και θα υπολογίσουμε το άθροισμα $\|x_1 + x_2 + \dots + x_n + x_{n+1}\|^2$. Είναι σαφές ότι $\langle x_1 + x_2 + \dots + x_n, x_{n+1} \rangle = 0$ συνεπώς έχουμε

$$\|x_1 + x_2 + \dots + x_n + x_{n+1}\|^2 = \|x_1 + x_2 + \dots + x_n\|^2 + \|x_{n+1}\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_n\|^2 + \|x_{n+1}\|^2.$$

□

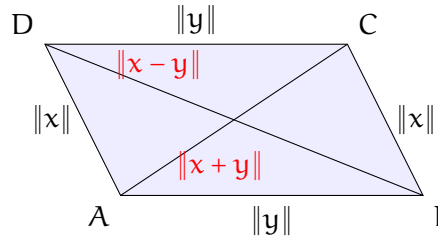
Πρόταση VII.4.7. Σε ένα χώρο με εσωτερικό γινόμενο $(X, \langle \cdot, \cdot \rangle)$ ισχύει η ισότητα του παραλληλ-
λογράμμου

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \text{ για κάθε } x, y \in X.$$

Αντιστρόφως αν σε ένα χώρο με νόρμα $(X, \| \cdot \|)$, ισχύει η παραπάνω ισότητα τότε η νόρμα προέρχεται από ένα εσωτερικό γινόμενο.

Απόδειξη. Προσθέτουμε κατά μέλη τις σχέσεις

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \\ \|x - y\|^2 &= \langle x - y, x - y \rangle = \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2 \end{aligned}$$



Σχήμα VII.2: Γεωμετρική αναπαράσταση της ισότητας του παραλληλογράμμου. Τα $\|x\|$, $\|y\|$ είναι τα μήκη των πλευρών και $\|x+y\|$, $\|x-y\|$ είναι τα μήκη των διαγωνίων.

οπότε προκύπτει η ισότητα του παραλληλογράμμου.

Αντιστρόφως, αν ισχύει η ισότητα του παραλληλογράμμου τότε το

$$\langle x, y \rangle := \frac{1}{4} (\|x+y\|^2 - \|x-y\|^2) \quad (\text{VII.11})$$

στην περίπτωση που ο διανυσματικός χώρος είναι πραγματικός και το

$$\langle x, y \rangle := \frac{1}{4} (\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2)$$

στην περίπτωση που ο διανυσματικός χώρος είναι μιγαδικός, ορίζουν εσωτερικό γινόμενο από το οποίο παράγεται η νόρμα $\|\cdot\|$.

Θα αποδείξουμε πρώτα την πραγματική περίπτωση. Υποθέτουμε ότι η νόρμα ικανοποιεί τον νόμο του παραλληλογράμμου, και θεωρούμε το εσωτερικό γινόμενο που ορίζεται από την εξίσωση (VII.11). Είναι σαφές ότι ικανοποιούνται οι ταυτότητες $\langle x, y \rangle = \langle y, x \rangle$ και $\|x\| = \sqrt{\langle x, x \rangle}$.

Παρατηρούμε ότι $\langle x+y+z, x+y+z \rangle = \langle x, z \rangle + \langle y, z \rangle$. Πράγματι, από τον κανόνα του παραλληλογράμμου έχουμε ότι

$$\|x+y+z\|^2 + \|x-y+z\|^2 = 2\|x+z\|^2 + 2\|y\|^2.$$

Αυτό μας δίνει ότι

$$\begin{aligned} \|x+y+z\|^2 &= 2\|x+z\|^2 + 2\|y\|^2 - \|x-y+z\|^2 \\ &= 2\|y+z\|^2 + 2\|x\|^2 - \|y-x+z\|^2, \end{aligned}$$

όπου η δεύτερη γραμμή προκύπτει εναλλάσσοντας τους ρόλους των x, y . Συνεπώς

$$\|x+y+z\|^2 = \|x\|^2 + \|y\|^2 + \|x+z\|^2 + \|y+z\|^2 - \frac{1}{2}\|x-y+z\|^2 - \frac{1}{2}\|y-x+z\|^2.$$

Με αντικατάσταση του z με $-z$ στην τελευταία εξίσωση παίρνουμε

$$\|x+y-z\|^2 = \|x\|^2 + \|y\|^2 + \|x-z\|^2 + \|y-z\|^2 - \frac{1}{2}\|x-y-z\|^2 - \frac{1}{2}\|y-x-z\|^2.$$

Τέλος χρησιμοποιώντας τα παραπάνω και τον ορισμό έχουμε

$$\begin{aligned} \langle x+y, z \rangle &= \frac{1}{4} (\|x+y+z\|^2 - \|x+y-z\|^2) \\ &= \frac{1}{4} (\|x+z\|^2 - \|x-z\|^2) + \frac{1}{4} (\|y+z\|^2 - \|y-z\|^2) \\ &= \langle x, z \rangle + \langle y, z \rangle \end{aligned}$$

Τώρα θα αποδείξουμε ότι $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$. Παρατηρούμε ότι η ζητούμενη ιδιότητα ισχύει για $\lambda = -1$, ενώ με βάση την προσθετική ιδιότητα και επαγωγή μπορούμε να αποδείξουμε ότι ισχύει

για κάθε $\lambda \in \mathbb{N}$ και αφού ισχύει και για $\lambda = -1$ ισχύει για κάθε $\lambda \in \mathbb{Z}$. Αν τώρα $\lambda = p/q$ με $p, q \in \mathbb{Z}, q \neq 0$ τότε για $x' = x/q$ έχουμε ότι

$$q\langle \lambda x, y \rangle = q\langle px', y \rangle = p\langle qx', y \rangle = p\langle x, y \rangle,$$

οπότε διαιρώντας με q έχουμε την ζητούμενη ιδιότητα:

$$\langle \lambda x, y \rangle = \lambda \langle x, y \rangle \text{ για κάθε } x, y \in V, \lambda \in \mathbb{Q}.$$

Για να μπορέσουμε να περάσουμε σε $\lambda \in \mathbb{R}$, παρατηρούμε ότι η συνάρτηση

$$\begin{aligned} f(t) : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ t &\longmapsto \frac{1}{t} \langle tx, y \rangle \end{aligned}$$

είναι συνεχής στο πεδίο ορισμού της, δηλαδή για κάθε ακολουθία (a_n) με $a_n \rightarrow l$ έχουμε $f(a_n) \rightarrow f(l)$. Επίσης παρατηρούμε ότι η $f(t) = \langle x, y \rangle$ για κάθε $t \in \mathbb{Q}$. Οπότε αν $\lambda \in \mathbb{R}^*$, διαλέγουμε μια ακολουθία ρητών αριθμών (a_n) με $a_n \rightarrow \lambda$ και έχουμε ότι $f(\lambda) = \frac{1}{\lambda} \langle \lambda x, y \rangle = \langle x, y \rangle$, δηλαδή το ζητούμενο. Η περίπτωση $\lambda = 0$ είναι προφανής.

Τέλος για την περίπτωση του μιγαδικού εσωτερικού γινομένου παρατηρούμε ότι $i\langle x, y \rangle = i\langle x, y \rangle$ και ότι $\langle x, y \rangle = \overline{\langle y, x \rangle}$. Για την απόδειξη της γραμμικότητας επαναλαμβάνουμε τα επιχειρήματα που εφραμόσαμε στο πραγματικό εσωτερικό γινόμενο για το πραγματικό $\operatorname{Re}\langle \cdot, \cdot \rangle$ και το φανταστικό $\operatorname{Im}\langle \cdot, \cdot \rangle$ μέρος του $\langle \cdot, \cdot \rangle$. $\square$

VII.4.1 Γωνία διανυσμάτων

Η ανισότητα Cauchy-Schwartz (VII.8) εξασφαλίζει ότι στην περίπτωση πραγματικού εσωτερικού γινομένου

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1$$

και συνεπώς μπορούμε να ορίσουμε την γωνία θ , $0 \leq \theta < 2\pi$ των διανυσμάτων x, y ώστε

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cos(\theta).$$

Στην περίπτωση που $\langle x, y \rangle = 0$ θα λέμε ότι τα x, y είναι κάθετα.

Ορισμός VII.4.8. Ένα σύνολο στοιχείων $S = \{e_\mu \in V, \mu \in I\}$ θα λέγεται ορθογώνιο αν και μόνο αν $\langle e_\mu, e_\nu \rangle = \delta_{\mu, \nu} \|e_\mu\|^2$. Θα λέγεται ορθοκανονικό αν επιπλέον $\|e_\mu\| = 1$ για κάθε $\mu \in I$.

Παρατήρηση VII.4.9. Ένα ορθογώνιο σύνολο S μπορεί να μετασχηματιστεί σε ορθοκανονικό $S' = \{e_\mu / \|e_\mu\| : \mu \in I\}$.

Πρόταση VII.4.10. Ένα ορθογώνιο σύνολο είναι γραμμικά ανεξάρτητο.

Απόδειξη. Πράγματι για οποιαδήποτε πεπερασμένη επιλογή $e_{i_1}, \dots, e_{i_r}$ από το ορθογώνιο σύνολο και σχέση

$$0_V = \sum_{v=1}^r \lambda_v e_{i_v}$$

έχουμε ότι

$$0_V = \langle 0_V, e_{i_j} \rangle = \left\langle \sum_{v=1}^r \lambda_v e_{i_v}, e_{i_j} \right\rangle = \sum_{v=1}^r \lambda_v \langle e_{i_v}, e_{i_j} \rangle = \lambda_j,$$

$\square$

δηλαδή όλοι οι συντελεστές που εμφανίζονται είναι μηδενικοί.

Πρόταση VII.4.11 (Ανάλυση Fourier). Αν σε ένα διανυσματικό χώρο έχουμε μια ορθοκανονική βάση $e_1, \dots, e_n$, τότε για κάθε διάνυσμα $v \in V$ γράφεται ως

$$\sum_{i=1}^n \langle v, e_i \rangle e_i.$$

Απόδειξη. Αφού τα $e_1, \dots, e_n$ είναι βάση, το τυχαίο στοιχείο $v \in V$ γράφεται ως

$$v = \sum_{\nu=1}^n \lambda_\nu e_\nu.$$

Υπολογίζουμε ότι

$$\langle v, e_i \rangle = \left\langle \sum_{\nu=1}^n \lambda_\nu e_\nu, e_i \right\rangle = \sum_{\nu=1}^n \lambda_\nu \langle e_\nu, e_i \rangle = \lambda_i.$$

□

Παράδειγμα VII.4.12. Θεωρούμε τον χώρο

$$X := \{f : [0, 2\pi] \rightarrow \mathbb{R} : f(0) = f(2\pi), f \text{ συνεχής} \}.$$

Ο χώρος αυτός μπορεί να ταυτιστεί με τον χώρο των περιοδικών συναρτήσεων $f : \mathbb{R} \rightarrow \mathbb{R}$ οι οποίες είναι περιοδικές με περίοδο 2π .

Στον χώρο αυτό εισάγουμε το εσωτερικό γινόμενο:

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt.$$

Παρατηρούμε ότι οι συναρτήσεις

$$e_0(s) = \frac{1}{\sqrt{2\pi}}, e_1(s) = \frac{1}{\sqrt{\pi}} \sin(s), e_2(s) = \frac{1}{\sqrt{\pi}} \cos(s), \dots, e_{2\nu-1}(s) = \frac{1}{\sqrt{\pi}} \sin(\nu s), e_{2\nu}(s) = \frac{1}{\sqrt{\pi}} \cos(\nu s), \dots$$

είναι στοιχεία του χώρου X , και μάλιστα ότι

$$\langle e_\nu, e_\mu \rangle = \delta_{\nu,\mu} \text{ για } \nu, \mu \in \mathbb{N}.$$

Πράγματι

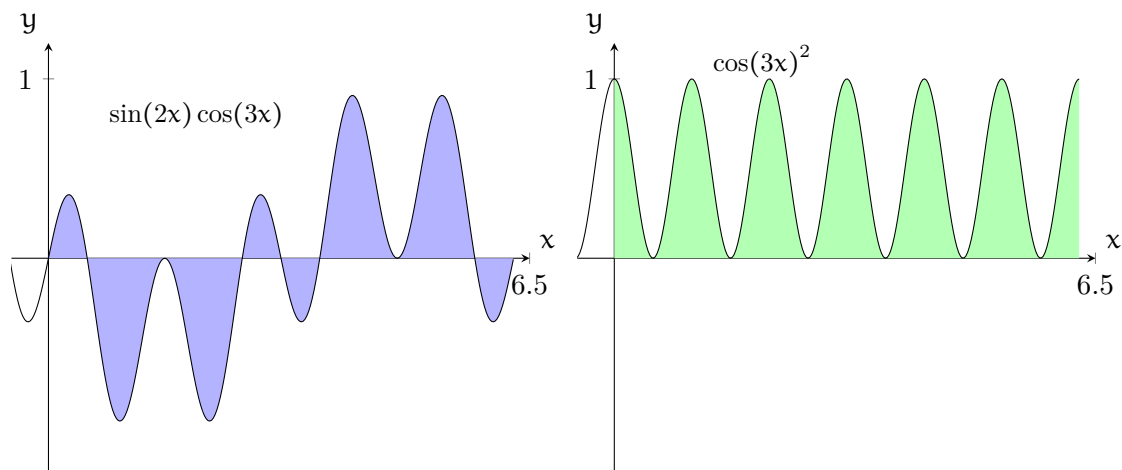
$$\begin{aligned} \int_0^{2\pi} \sin(\nu t) \sin(\mu t) dt &= \frac{1}{2} \int_0^{2\pi} (\cos((\nu - \mu)t) - \cos((\nu + \mu)t)) dt = \pi \cdot \delta_{\nu,\mu} \\ \int_0^{2\pi} \cos(\nu t) \cos(\mu t) dt &= \frac{1}{2} \int_0^{2\pi} (\cos((\nu - \mu)t) + \cos((\nu + \mu)t)) dt = \pi \cdot \delta_{\nu,\mu} \\ \int_0^{2\pi} \sin(\nu t) \cos(\mu t) dt &= \frac{1}{2} \int_0^{2\pi} (\sin((\nu - \mu)t) + \sin((\nu + \mu)t)) dt = 0 \\ \int_0^{2\pi} e_0(t) e_\mu(t) dt &= \delta_{0,\mu}. \end{aligned}$$

VII.5 Μια εφαρμογή στις διαφορικές εξισώσεις

VII.5.1 Διαφορικές εξισώσεις και εκθετικές συναρτήσεις

ΑΣ ξεκινήσουμε για προθέρμανση με το εξής πρόβλημα διαφορικών εξισώσεων: Να βρεθεί μια συνάρτηση $f(x)$ ώστε

$$f'(x) = \lambda f(x). \quad (\text{VII.12})$$



Σχήμα VII.3: Το αριστερό εμβαδόν είναι 0, τα εμβαδά κάτω από τον άξονα των x είναι αρνητικά. Το δεξί εμβαδόν είναι π

Αν υποθέσουμε ότι έχει αναλυτική λύση

$$f(x) = \sum_{i=0}^{\infty} a_i x^i$$

τότε η εξίσωση (VII.12) μας δίνει ότι

$$\sum_{i=0}^{\infty} a_i i x^{i-1} = \lambda \sum_{i=0}^{\infty} a_i x^i$$

από όπου προκύπτει ότι

$$a_{i+1} = a_i \lambda / i, \text{ συνεπώς } a_i = \frac{\lambda^i}{i!} a_0.$$

και καταλήγουμε στην

$$f(x) = a_0 \sum_{i=0}^{\infty} \frac{(\lambda x)^i}{i!} = a_0 e^{\lambda x}.$$

VII.5.2 Διαφορικές εξισώσεις n -τάξης

Θεωρούμε την διαφορική εξίσωση

$$\frac{d^n f}{dt^n} + a_{n-1} \frac{d^{n-1} f}{dt^{n-1}} + \dots + a_1 \frac{df}{dt} + a_0 f = 0. \quad (\text{VII.13})$$

Δηλαδή θεωρούμε την γραμμική συνάρτηση

$$D : C^n(\mathbb{R}, \mathbb{R}) \longrightarrow C^n(\mathbb{R}, \mathbb{R})$$

$$f(t) \longmapsto \frac{df}{dt}$$

από τον χώρο $C^n(\mathbb{R}, \mathbb{R})$ των n -φορές παραγωγίσιμων συναρτήσεων στον ευατό του. Στην συνέχεια, για το πολυώνυμο

$$p(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0,$$

θεωρούμε την γραμμική συνάρτηση

$$p(D) : C^n(\mathbb{R}, \mathbb{R}) \longrightarrow C^n(\mathbb{R}, \mathbb{R})$$

$$f(t) \longmapsto p(D)f = p\left(\frac{d}{dt}\right)f(t)$$

και αναζητούμε τον πυρήνα $V = \ker p(D)$.

Παρατήρηση VII.5.1. Επειδή στην συνέχεια θα αναζητήσουμε ρίζες του πολυωνύμου $p(t)$ είναι βολικότερο να μεταφερθούμε στον χώρο $C^n(\mathbb{R}, \mathbb{C})$, δηλαδή να πάρουμε n -φορές παραγωγίσιμες συναρτήσεις με τιμές στους μιγαδικούς, δείτε και το παράδειγμα VII.5.3.

Αν για έχουμε

$$p(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2}\dots(x - \lambda_k)^{r_k}$$

την ανάλυση του $p(x)$ σε γινόμενο διαφορετικών ριζών και

$$W_j = \ker(D - \lambda_j \text{Id})^{r_j}$$

τότε

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_k.$$

Λήμμα VII.5.2. Αν $f \in C^n(\mathbb{R}, \mathbb{R})$ τότε για $r \in \mathbb{N}$ ισχύει

$$(D - \lambda \text{Id})^r = e^{\lambda t} D^r (e^{-\lambda t} f) = e^{\lambda t}.$$

Απόδειξη. Αν $r = 0$ το αποτέλεσμα είναι προφανές. Υποθέτουμε ότι το αποτέλεσμα ισχύει για $r = n$, υπολογίζουμε ότι

$$(D - \lambda \text{Id})^{n+1}f(t) = (D - \lambda \text{Id})(D - \lambda \text{Id})^n f = (D - \lambda \text{Id})e^{\lambda t} D^n (e^{-\lambda t} f) = e^{\lambda t} D^{n+1} (e^{-\lambda t} f).$$

□

Έχουμε ότι $(D - \lambda \text{Id})^r f = 0$ αν και μόνο αν $D^r (e^{-\lambda t} f) = 0$. Παρατηρούμε ότι $D^r g(t) = 0$ αν και μόνο αν $g(t) = b_0 + b_1 t + \dots + b_{r-1} t^{r-1}$. Συνεπώς

$$f(t) = e^{\lambda t} (b_0 + b_1 t + \dots + b_{r-1} t^{r-1}).$$

Δηλαδή οι συναρτήσεις $e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_i t} t^{r_i-1}$ παράγουν τον χώρο W_i και επειδή είναι γραμμικά ανεξάρτητες αποτελούν μία βάση του.

Παράδειγμα VII.5.3. Θεωρούμε την διαφορική εξίσωση

$$\frac{d^2 f(t)}{dt} = -af(t), \text{ με } a > 0$$

η οποία σχετίζεται με τον αρμονικό ταλαντωτή στην φυσική και μπορεί να αντιμετωπιστεί και καθαρά με αναλυτικές μεθόδους, δείτε [22, παρ. 4.2a]. Το πολυώνυμο $p(x) = x^2 + a$ έχει δύο μιγαδικές ρίζες τις $\pm i\sqrt{a}$. Οι λύσεις της διαφορικής εξίσωσης είναι της μορφής

$$f(t) = \lambda_1 e^{i\sqrt{a}t} + \lambda_2 e^{-i\sqrt{a}t},$$

όπου τα $\lambda_1, \lambda_2 \in \mathbb{C}$ προσδιορίζονται από τις αρχικές συνθήκες $f(0), f'(0)$. Το ότι χρειαζόμαστε δύο αρχικές συνθήκες σχετίζεται με την γραφή της εξίσωσης ως σύστημα διαφορικών εξισώσεων, δείτε

και την παρατήρηση [VII.5.11](#). Μέσα σε αυτές τις λύσεις μπορούμε να αναζητήσουμε λύσεις στο $C^n(\mathbb{R}, \mathbb{R})$, αφού

$$\begin{aligned}\lambda_1 e^{i\sqrt{a}t} + \lambda_2 e^{-i\sqrt{a}t} &= \lambda_1 (\cos(\sqrt{a}t) + i \sin(\sqrt{a}t)) + \lambda_2 (\cos(-\sqrt{a}t) + i \sin(-\sqrt{a}t)) \\ &= (\lambda_1 + \lambda_2) \cos(\sqrt{a}t) + i(\lambda_1 - \lambda_2) \sin(\sqrt{a}t)\end{aligned}$$

Η τελευταία λύση είναι πραγματική αν $\operatorname{Re}(\lambda_1) = \operatorname{Re}(\lambda_2)$ και $\operatorname{Im}(\lambda_1) = -\operatorname{Im}(\lambda_2)$, οπότε καταλήγουμε στην πραγματική οικογένεια συναρτήσεων:

$$\mu_1 \cos(\sqrt{a}t) + \mu_2 \sin(\sqrt{a}t)$$

με $\mu_1 = 2\operatorname{Re}(\lambda_1)$ και $\mu_2 = -2\operatorname{Im}(\lambda_1)$.

VII.5.3 Ο εκθετικός πίνακας

Έχουμε ήδη δει ότι έχει νόημα να ορίσουμε για $A \in \mathbb{F}^{n,n}$ και $f(x) = \sum_{i=0}^n a_i x^i \in \mathbb{F}[x]$ τον πίνακα

$$f(A) = \sum_{i=0}^n a_i A^i \in \mathbb{F}^{n,n}.$$

Θέλουμε να επεκτείνουμε τον παραπάνω ορισμό για συναρτήσεις οι οποίες εκφράζονται μέσω δυναμοσειρών όπως τις ορίσαμε στην παράγραφο [I.8.8](#). Για παράδειγμα ιδιαίτερα ενδιαφέρον είναι ο εκθετικός πίνακας

$$\exp(A) := \sum_{i=0}^{\infty} \frac{A^i}{i!}.$$

Για ένα τέτοιο άπειρο άθροισμα πινάκων χρειαζόμαστε την έννοια της σύγκλισης. Θα χρειαστεί να εισάγουμε μερικές έννοιες για νόρμες πινάκων. Ο χώρος των $n \times n$ πινάκων είναι ένας διανυσματικός χώρος, διάστασης n^2 οπότε μπορούμε να ορίσουμε την έννοια της νόρμας όπως κάναμε στην παράγραφο [VII.1](#). Χρειαζόμαστε νόρμες οι οποίες να σέβονται την πράξη του πολλαπλασιασμού πινάκων, να είναι δηλαδή υποπολλαπλασιαστικές όπως ορίστηκαν στο [VII.2.2](#).

Λήμμα VII.5.4. *Αν μια σειρά πινάκων ικανοποιεί*

$$\left\| \sum_{n=0}^{\infty} a_n A^n \right\| \leq \sum_{n=0}^{\infty} M_n < \infty$$

τότε αυτή συγκλίνει.

Απόδειξη. Ο χώρος των $n \times n$ είναι πεπερασμένης διάστασης πάνω από το πλήρες σώμα $\mathbb{F}$ συνεπώς είναι πλήρης. Αυτό σημαίνει ότι μια ακολουθία είναι Cauchy αν και μόνο αν συγκλίνει. Η συνθήκη Cauchy για μια νόρμα πινάκων είναι η

$$\text{Για κάθε } \epsilon > 0 \text{ υπάρχει } n_0 \in \mathbb{N}, \text{ ώστε } n \geq m > n_0 \text{ ισχύει } \left\| \sum_{i=m}^n a_i A^i \right\| < \epsilon$$

□

Ορισμός VII.5.5. Για σταθερό $A \in \mathbb{F}^{n,n}$ ορίζουμε την εκθετική συνάρτηση $\mathbb{F} \rightarrow \mathbb{F}^{n,n}$ με

$$\exp(Ax) = \sum_{n=0}^{\infty} \frac{A^n x^n}{n!}.$$

Για να δείξουμε ότι η παραπάνω συνάρτηση είναι καλά ορισμένη αρκεί να χρησιμοποιήσουμε το λήμμα VII.5.4 και να παρατηρήσουμε ότι για μια υποπολλαπλασιαστική νόρμα έχουμε

$$\left\| \sum_{n=0}^{\infty} \frac{A^n x^n}{n!} \right\| \leq \sum_{n=0}^{\infty} \frac{\|A\|^n |x|^n}{n!}$$

το οποίο συγκλίνει αφού αναπαριστά το $e^{\|A\||x|}$.

Πρόταση VII.5.6. Στην περίπτωση που ο πίνακας A είναι διαγωνοποιήσιμος, δηλαδή υπάρχει αντιστρέψιμος πίνακας Q ώστε $A = Q \operatorname{diag}(\lambda_1, \dots, \lambda_n) Q^{-1}$ έχουμε

$$\exp(At) = Q \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) Q^{-1}.$$

Απόδειξη. Από τον ορισμό έχουμε

$$\exp(At) = \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} = Q \operatorname{diag} \left(\sum_{i=0}^{\infty} \frac{\lambda_1^i t^i}{i!}, \dots, \sum_{i=0}^{\infty} \frac{\lambda_n^i t^i}{i!} \right) Q^{-1} = Q \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) Q^{-1}.$$

□

Πρόταση VII.5.7. Έχουμε ότι η παράγωγος

$$\frac{d}{dt} \exp(At) = A \cdot \exp(At).$$

Απόδειξη.

$$\frac{d}{dt} \exp(At) = \frac{d}{dt} \left(\sum_{i=0}^{\infty} \frac{A^i t^i}{i!} \right) = \sum_{i=1}^{\infty} \frac{A^i i t^{i-1}}{i!} = A \sum_{i=1}^{\infty} \frac{A^{i-1} t^{i-1}}{(i-1)!} = A \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} = A \cdot \exp(At).$$

□

Πρόταση VII.5.8. Για πίνακες $A, B \in \mathbb{C}$ με $AB = BA$ ισχύει

$$\exp((A+B)t) = \exp(At) \exp(Bt).$$

Απόδειξη. Επειδή $AB = BA$, έχουμε ότι για $k \in \mathbb{N}$

$$(A+B)^v = \sum_{\nu=0}^v \binom{v}{\nu} A^\nu B^{v-\nu}.$$

Στην συνέχεια υπολογίζουμε

$$\begin{aligned} \exp((A+B)t) &= \sum_{i=0}^{\infty} \frac{(A+B)^i t^i}{i!} = \sum_{i=0}^{\infty} \sum_{\nu=0}^i \binom{i}{\nu} A^\nu B^{i-\nu} \frac{t^i}{i!} = \sum_{i=0}^{\infty} \sum_{\nu=0}^i \frac{A^\nu t^\nu}{\nu!} \frac{B^{i-\nu} t^{i-\nu}}{(i-\nu)!} \\ &= \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} \sum_{i=0}^{\infty} \frac{B^i t^i}{i!} = \exp(At) \exp(Bt). \end{aligned}$$

□

Πρόταση VII.5.9. Για κάθε πίνακα $A \in \mathbb{C}^{n,n}$ ο εκθετικός πίνακας $\exp(A)$ είναι αντιστρέψιμος. Το $\exp(At)$ είναι μια οικογένεια αντιστρέψιμων πινάκων η οποία για $t = 0$ δίνει:

$$\exp(At)|_{t=0} = \mathbb{I}_n, \quad \frac{d}{dt} \exp(At)|_{t=0} = A.$$

Απόδειξη. Το ότι $\exp(At)|_{t=0} = \exp(\mathbf{0}_{\mathbb{C}^{n,n}}) = \mathbb{I}_n$ προκύπτει από τον ορισμό του εκθετικού πίνακα **VII.5.5**. Η αντιστρεψιμότητα προκύπτει από το ότι

$$\mathbb{I}_n = \mathbf{0}_{\mathbb{C}^{n,n}} = \exp(At - At) = \exp(At) \exp(-At).$$

Ειδικότερα έχουμε ότι

$$\exp(At)^{-1} = \exp(-At).$$

□

Πρόταση VII.5.10. Στην περίπτωση που ο A είναι ένα Jordan μπλοκ

$$A = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

Απόδειξη. Στην περίπτωση που ο A είναι ένα Jordan μπλοκ

$$A = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} = \text{diag}(\lambda, \dots, \lambda) + \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

Θέτουμε D για το διαγώνιο κομμάτι και N για το μηδενοδύναμο και έχουμε $DN = ND$ οπότε

$$\exp(At) = \exp((D + N)t) = \exp(Dt) \exp(Nt).$$

Για τις δυνάμεις του $N \in \mathbb{R}^{n,n}$ έχουμε

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ \vdots & & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}, \dots, N^{n-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, N^n = \mathbf{0}$$

οπότε υπολογίζουμε ότι

$$\exp(Nt) = \sum_{i=0}^{n-1} \frac{N^i}{i!} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & \ddots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

και τελικά καταλήγουμε στο

$$\exp(At) = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & \ddots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

□

VII.5.4 Συστήματα διαφορικών εξισώσεων

Ας θεωρήσουμε δύο συναρτήσεις $u(t), v(t)$ οι οποίες ικανοποιούν την σχέση

$$\begin{aligned} \frac{du}{dt} &= a_{11}u(t) + a_{12}v(t) \\ \frac{dv}{dt} &= a_{21}u(t) + a_{22}v(t) \end{aligned}$$

με $u(0) = u_0, v(0) = v_0$ και $a_{ij}, u_0, v_0 \in \mathbb{R}$. Το πρόβλημα είναι να βρεθούν οι συναρτήσεις $u(t), v(t)$. Γενικότερα μπορούμε να έχουμε ένα διάνυσμα συναρτήσεων $F(t) = (f_1(t), \dots, f_n(t))^t$ και μία σχέση

$$\frac{d}{dt}F(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix} = AF(t) = A \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}, \quad (\text{VII.14})$$

όπου $A \in \mathbb{R}^{n,n}$ και $F(0) = u_0 \in \mathbb{R}^n$. Από την πρόταση [VII.5.7](#) έχουμε ότι η

$$F(t) = \exp(At)u_0,$$

είναι μια λύση της διαφορικής εξίσωσης, η οποία για $t = 0$ εξασφαλίζει ότι ικανοποιείται και η αρχική συνθήκη $F(0)$. Το πρόβλημα της επίλυσης διαφορικών εξισώσεων αυτού του τύπου ανάγεται στο πρόβλημα του υπολογισμού του εκθετικού πίνακα $\exp(At)$. Αν ο πίνακας A είναι διαγωνοποιήσιμος, δηλαδή υπάρχει Q αντιστρέψιμος ώστε $A = Q \text{diag}(\lambda_1, \dots, \lambda_n) Q^{-1}$, τότε σύμφωνα με την πρόταση [VII.5.6](#) έχουμε ότι

$$F(t) = Q \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) Q u_0.$$

είναι η ζητούμενη λύση της διαφορικής εξίσωσης.

Στην περίπτωση που ο πίνακας έχει μια κανονική μορφή Jordan

$$Q^{-1}AQ = \begin{pmatrix} J(\lambda_1, k_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J(k_r, \lambda_r) \end{pmatrix}$$

τότε

$$\exp(At) = Q \begin{pmatrix} e^{\lambda_1 t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{k_1-1}}{(k_1-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & \ddots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \\ \vdots & \ddots & \ddots & 0 & \\ 0 & \cdots & \cdots & 0 & e^{\lambda_r t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{k_r-1}}{(k_r-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & \ddots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \end{pmatrix} Q^{-1}$$

και η λύση του συστήματος διαφορικών εξισώσεων δίνεται από το $F(t) = Q \exp(At) Q^{-1} u_0$.

Παρατήρηση VII.5.11. Η διαφορική εξίσωση (VII.13) n -τάξης της παραγράφου VII.5.2 μπορεί να μετατραπεί σε ένα σύστημα διαφορικών εξισώσεων θέτοντας:

$$f_1(t) = \frac{d}{dt} f(t), f_2(t) = \frac{d}{dt} f_1(t), \dots, f_{n-1}(t) = \frac{d}{dt} f_{n-2}(t),$$

οπότε η (VII.13) είναι ισοδύναμη με το σύστημα:

$$\frac{d}{dt} \begin{pmatrix} f(t) \\ f_1(t) \\ f_2(t) \\ \vdots \\ f_{n-1}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{pmatrix} \begin{pmatrix} f(t) \\ f_1(t) \\ f_2(t) \\ \vdots \\ f_{n-1}(t) \end{pmatrix}$$

Το χαρακτηριστικό πολυώνυμο του πίνακα του συστήματος είναι το $p(x)$ και η ανάλυση Jordan του πίνακα ταιριάζει με την λύση της παραγράφου VII.5.2.

VII.5.5 Πορτρέτα Φάσης, δυναμικά συστήματα

Θέλουμε να μελετήσουμε την συμπεριφορά της λύσης VII.14

$$F(t) : \mathbb{R} \longrightarrow \mathbb{R}^n.$$

Μπορούμε να σχεδιάσουμε την $F(t)$ ως μια παραμετρική καμπύλη στον χώρο $\mathbb{R}^n$, και την βλέπουμε σαν την εξέλιξη ενός φαινομένου στον χώρο $\mathbb{R}^n$.

Ορισμός VII.5.12. Την καμπύλη $F \exp(At) u_0$ θα την λέμε τροχιά του σημείου u_0 .

Το κλειδί για την συμπεριφορά της $F(t)$ είναι οι ιδιοτιμές του πίνακα A .

Παρατηρούμε ότι αν $\lambda = a + ib \in \mathbb{C}$, $a, b \in \mathbb{R}$, τότε

$$e^{\lambda t} = e^{(a+ib)t} = e^{at} e^{ibt} = e^{at} (\cos(bt) + i \sin(bt)),$$

δηλαδή $|e^{\lambda t}| = e^{at}$ και η ασυμπτωτική συμπεριφορά της $e^{\lambda t}$ καθορίζεται από το $\operatorname{Re} \lambda$.

Ορισμός VII.5.13. Η διαφορική εξίσωση (VII.14) θα λέγεται

- Ευσταθής, και σε αυτή την περίπτωση $\lim_{t \rightarrow \infty} e^{At} = 0$ αν τα πραγματικά μέρη των ιδιοτιμών λ_i του πίνακα A ικανοποιούν $\operatorname{Re} \lambda_i < 0$.
- Ουδέτερα ευσταθής, αν ισχύει ότι $\operatorname{Re} \lambda_i \leq 0$ και κάποιο $\operatorname{Re} \lambda_i = 0$
- Ασταθής αν κάποια ιδιοτιμή έχει $\operatorname{Re} \lambda_i > 0$.

Παρατήρηση VII.5.14. Η διαφορική εξίσωση (VII.14) μαζί με την γνώση της αρχικής συνθήκης $F(0) = u_0 \in \mathbb{R}^n$ καθορίζει την εξέλιξη της $F(t) = \exp(At)u_0$ στο μέλλον ($t > 0$) και στο παρελθόν ($t < 0$). Από κάθε σημείο $u \in \mathbb{R}^n$ υπάρχει μια μοναδική τροχιά η οποία περνάει από αυτό και προσδιορίζεται από την παραπάνω λύση της διαφορικής εξίσωσης. Αυτή η παρατήρηση μας δίνει την δυνατότητα να ορίσουμε μια δράση της ομάδας $(\mathbb{R}, +)$ πάνω στο $\mathbb{R}^n$ (όπως αυτή την ορίσαμε στην παράγραφο I.8.5).

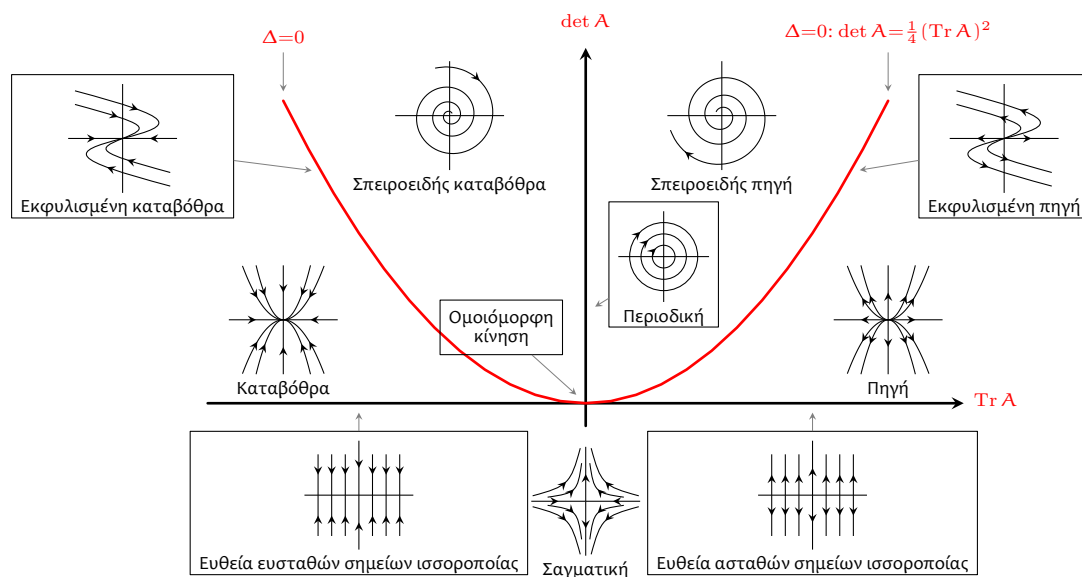
$$\begin{aligned} \mathbb{R} \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (t, u) &\longmapsto t \cdot u = \exp\{At\}u. \end{aligned}$$

η οποία είναι μια δράση ομάδας δηλαδή

- $0 \cdot u = \exp(A0)u = \mathbb{I}_n u = u$
- $(t_1 + t_2) \cdot u = \exp(A(t_1 + t_2))u = \exp(At_1) \exp(At_2)u = t_1 \cdot (t_2 \cdot u)$.

Η ύπαρξη αυτής της δράσης σχετίζεται με την αναπαράσταση την οποία θα αναφέρουμε στην (IX.2).

Στην περίπτωση $n = 2$, έχουμε ζωγραφίσει την μορφή των λύσεων στο σχήμα VII.4



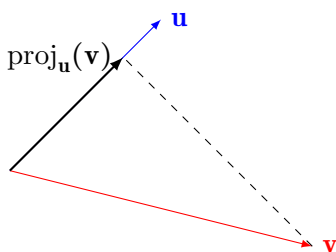
Σχήμα VII.4: Ταξινόμηση των πορτρέτων φάσης στο επίπεδο $(\det A, \operatorname{tr} A)$

VII.6 Ορθοκανονικοποίηση

Στην παράγραφο αυτή θα περιγράψουμε μία μέθοδο η οποία δοσμένου ενός πεπερασμένου συνόλου γραμμικά ανεξάρτητων διανυσμάτων $v_1, \dots, v_n$ σε ένα χώρο με εσωτερικό γινόμενο δίνει μια ορθοκανονική βάση του χώρου $\langle v_1, \dots, v_n \rangle$ που αυτά παράγουν.

Ορισμός VII.6.1. Η προβολή ενός στοιχείου $v \in V$ επί ενός μη μηδενικού στοιχείου $u \in V$ ορίζεται ως

$$\text{proj}_u(v) = \frac{\langle v, u \rangle}{\langle u, u \rangle} u.$$



Σχήμα VII.5: Γεωμετρική αναπαράσταση της προβολής του διανύσματος v στο διάνυσμα u .

Παρατήρηση VII.6.2. Είναι σαφές ότι το διάνυσμα $\text{proj}_u(v)$ είναι παράλληλο στο u και ότι το $v - \text{proj}_u(v)$ είναι κάθετο με το u αφού

$$\langle v - \text{proj}_u(v), u \rangle = \langle v, u \rangle - \frac{\langle v, u \rangle}{\langle u, u \rangle} \langle u, u \rangle = 0.$$

Η διαδικασία ορθοκανονικοποίησης κατά Gram-Schmidt είναι η παρακάτω:

Μέθοδος 5 (Gram-Schmidt ορθοκανονικοποίηση). Δίνεται ένα πεπερασμένο σύνολο διανυσμάτων $v_1, \dots, v_n$ για να βρούμε μια ορθοκανονική βάση $e_1, \dots, e_n$ του χώρου που παράγουν εκτελούμε τα παρακάτω βήματα:

- 1 Θέτουμε $u_1 = v_1$
- 2 Θέτουμε $u_2 = v_2 - \text{proj}_{u_1}(v_2)$
- 3 Θέτουμε $u_3 = v_3 - \text{proj}_{u_1}(v_3) - \text{proj}_{u_2}(v_3)$
- 4 Θέτουμε $u_4 = v_4 - \text{proj}_{u_1}(v_4) - \text{proj}_{u_2}(v_4) - \text{proj}_{u_3}(v_4)$
- 5 $\vdots$
- 6 Θέτουμε $u_n = v_n - \sum_{j=1}^{n-1} \text{proj}_{u_j}(v_n)$.
- 7 Θέτουμε $e_i = \frac{u_i}{\|u_i\|}$.

Πρόταση VII.6.3. Η μέθοδος Gram-Schmidt δίνει πράγματι μια ορθοκανονική βάση του χώρου που παράγουν τα $v_1, \dots, v_n$.

Απόδειξη. Πράγματι, παρατηρούμε ότι το σύνολο των διανυσμάτων $u_1, \dots, u_n$ είναι ανά δύο κάθετα. Αυτό για $n = 1$ δεν έχει νόημα και για $n = 2$ το είδαμε στην παρατήρηση VII.6.2. Υποθέτουμε επαγωγικά ότι τα $u_1, \dots, u_k$ είναι ανά δύο κάθετα και θα δείξουμε ότι και το u_{k+1} είναι κάθετο με όλα τα προηγούμενα. Έχουμε ότι για $1 \leq i \leq k$

$$\begin{aligned}\langle u_{k+1}, u_i \rangle &= \left\langle v_{k+1} - \sum_{j=1}^k \text{proj}_{u_j}(v_{k+1}), u_i \right\rangle \\ &= \langle v_{k+1}, u_i \rangle - \langle \text{proj}_{u_i}(v_{k+1}), u_i \rangle \\ &= \langle v_{k+1}, u_i \rangle - \frac{\langle v_{k+1}, u_i \rangle}{\langle u_i, u_i \rangle} \langle u_i, u_i \rangle = 0.\end{aligned}$$

Παρατηρούμε ότι τα $e_i \in \langle v_1, \dots, v_n \rangle$ για $1 \leq i \leq n$, συνεπώς $\langle e_1, \dots, e_n \rangle \subset \langle v_1, \dots, v_n \rangle$. Τέλος τα $\{e_1, \dots, e_n\}$ είναι ως ορθοκανονικό σύνολο γραμμικά ανεξάρτητο άρα $\langle e_1, \dots, e_n \rangle = \langle v_1, \dots, v_n \rangle$. $\square$

Παράδειγμα VII.6.4. Έχουμε τα διανύσματα

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

από τα οποία θα κατασκευάσουμε μια ορθοκανονική βάση με την μέθοδο Gram-Schmidt. Ξεκινάμε με

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Υπολογίζουμε

$$\text{proj}_{u_1}(v_2) = \frac{\langle v_2, u_1 \rangle}{\|u_1\|^2} u_1 = \frac{2}{7} u_1$$

Στη συνέχεια, υπολογίζουμε το u_2 ως εξής:

$$u_2 = v_2 - \text{proj}_{u_1}(v_2) = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{2}{7} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 12 \\ 3 \\ -6 \end{pmatrix}$$

Συνεχίζουμε με

$$\begin{aligned}\text{proj}_{u_1}(v_3) &= \frac{\langle v_3, u_1 \rangle}{\|u_1\|^2} u_1 = \frac{7}{14} u_1 = \frac{1}{2} u_1 \\ \text{proj}_{u_2}(v_3) &= \frac{\langle v_3, u_2 \rangle}{\|u_2\|^2} u_2 = \frac{-21}{27} u_2\end{aligned}$$

και

$$\begin{aligned}u_3 &= v_3 - \text{proj}_{u_1}(v_3) - \text{proj}_{u_2}(v_3) \\ &= \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \frac{3}{27} \begin{pmatrix} 12 \\ 3 \\ -6 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/3 \\ -1/6 \end{pmatrix}.\end{aligned}$$

Τέλος υπολογίζουμε ότι

$$\|u_1\| = \sqrt{14}, \quad \|u_2\| = \sqrt{27/7}, \quad \|u_3\| = \sqrt{1/6}$$

και έχουμε την ορθοκανονική βάση

$$e_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{14}}u_1, \quad e_2 = \frac{u_2}{\|u_2\|} = \sqrt{7/27}u_2, \quad e_3 = \frac{u_3}{\|u_3\|} = \sqrt{6}u_3.$$

VII.6.1 Εφαρμογή σε προβλήματα βελτιστοποίησης

Ορισμός VII.6.5. Θεωρούμε ένα διανυσματικό χώρο $(V, \|\cdot\|)$ με νόρμα και ένα υποσύνολο $\emptyset \neq \Sigma \subset V$. Ένα $x \in \Sigma$ θα λέγεται βέλτιστη προσέγγιση του $z \in V$ αν ισχύει

$$\|x - z\| \leq \|x - y\| \text{ για κάθε } y \in \Sigma.$$

Παρατήρηση VII.6.6. Ένα τέτοιο στοιχείο μπορεί να μην υπάρχει, θεωρήστε ένα ανοιχτό διάστημα $\Sigma = (0, 1) \in \mathbb{R}$, ή να μην είναι μοναδικό.

Έστω $(V, \langle \cdot, \cdot \rangle)$ ένας χώρος με εσωτερικό γινόμενο και $W \subset V$ ένας υπόχωρος του. Θεωρούμε μια ορθοκανονική βάση $\{w_1, \dots, w_n\}$ του υποχώρου W και ορίζουμε το στοιχείο

$$x = \langle z, w_1 \rangle w_1 + \dots + \langle z, w_n \rangle w_n. \quad (\text{VII.15})$$

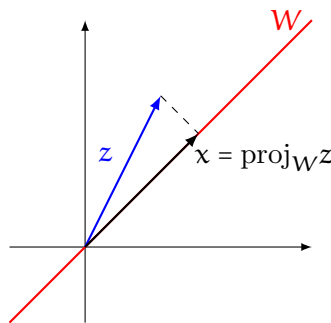
Για κάθε $1 \leq j \leq n$ υπολογίζουμε

$$\langle z - x, w_j \rangle = \langle z, w_j \rangle - \langle x, w_j \rangle = \langle z, w_j \rangle - \langle z, w_j \rangle = 0,$$

λαμβάνοντας υπόψιν ότι $\langle w_i, w_j \rangle = \delta_{ij}$. Συνεπώς το διάνυσμα $z - x$ είναι ορθογώνιο με κάθε γραμμικό συνδυασμό των $w_1, \dots, w_n$ δηλαδή με κάθε στοιχείο του υποχώρου W :

$$\langle z - x, w \rangle = 0 \text{ για κάθε } w \in W. \quad (\text{VII.16})$$

Το σημείο x είναι η ορθογώνια προβολή του z στον χώρο W . Θα αποδείξουμε ότι το x είναι



Σχήμα VII.6: Προβολή του z στον υπόχωρο W , η διακεκομμένη γραμμή είναι η ελάχιστη απόσταση του z από τον υπόχωρο W .

πραγματικά μια βέλτιστη προσέγγιση του z από στοιχεία του W . Θεωρούμε ένα στοιχείο $y \in W$. Ισχύει ότι $x - y \in W$ και η (VII.16) δίνει $\langle z - x, x - y \rangle = 0$. Σύμφωνα με το Πυθαγόρειο θεώρημα έχουμε ότι

$$\|z - y\|^2 = \|(z - x) + (x - y)\|^2 = \|z - x\|^2 + \|x - y\|^2 \geq \|z - x\|^2,$$

δηλαδή το x είναι μια βέλτιστη προσέγγιση. Αν $x \neq y$ τότε $\|z - y\| > \|z - x\|$, δηλαδή η βέλτιστη προσέγγιση είναι μοναδική.

Παράδειγμα VII.6.7. Θεωρούμε τον χώρο

$$X := \{f : [0, 2\pi] \rightarrow \mathbb{R} : f(0) = f(2\pi), f \text{ συνεχής} \}.$$

του παραδείγματος VII.6.7 και τον πεπερασμένο υπόχωρο

$$T_n = \left\{ x \in X : x(t) = \alpha_0 + \sum_{\nu=1}^n (\alpha_\nu \cos(\nu t) + \beta_\nu \sin(\nu t)) \right\}$$

των τριγωνομετρικών πολυώνυμων βαθμού το πολύ n . Οι συναρτήσεις

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, e_{2\nu-1}(t) = \frac{1}{\sqrt{\pi}} \sin(\nu t), e_{2\nu}(t) = \frac{1}{\sqrt{\pi}} \cos(\nu t), \nu = 1, \dots, n$$

αποτελούν μια ορθοκανονική βάση του T_n . Η βέλτιστη προσέγγιση μιας τυχαίας περιοδικής συνάρτησης από τον T_n είναι το μερικό άθροισμα της λεγόμενης σειράς Fourier της x :

$$S_n := \frac{1}{2\pi} \int_0^{2\pi} x(t) dt + \frac{1}{\sqrt{\pi}} \sum_{\nu=1}^{2n} \left(\int_0^{2\pi} x(t) e_\nu(t) dt \right) e_\nu.$$

Η μελέτη των σειρών Fourier είναι κλάδος της Ανάλυσης γνωστός ως Αρμονική Ανάλυση και έχει πολλές εφαρμογές στα Θεωρητικά και Εφαρμοσμένα Μαθηματικά, στην Φυσική και στην τεχνολογία.

VII.6.2 Ορθογώνιες προβολές

Έστω ένας πεπερασμένης διάστασης διανυσματικός χώρος V με εσωτερικό γινόμενο και έστω W ένας υπόχωρος. Η μέθοδος Gram-Schmidt μας δίνει μία μέθοδο να κατασκευάσουμε μια ορθοκανονική βάση $B_W = \{e_1, \dots, e_n\}$ του W την οποία μπορούμε να την επεκτείνουμε σε μία ορθοκανονική βάση $B_V = \{e_1, \dots, e_n, e_{n+1}, \dots, e_m\}$ του V . Είναι σαφές ότι αν θέσουμε

$$W^\perp = \langle e_{n+1}, \dots, e_m \rangle$$

τότε

$$V = W \oplus W^\perp.$$

Μπορούμε να εφοδιάσουμε τον χώρο των γραμμικών συναρτήσεων $\text{End}(V)$ από τον $V \rightarrow V$ με δύο στοιχεία $p_W, p_{W^\perp}$, τις ορθογώνιες προβολές, όπου

$$p_W : \begin{cases} V \longrightarrow W \subset V \\ w + w' = v \longmapsto w \end{cases} \quad p_{W^\perp} : \begin{cases} V \longrightarrow W^\perp \subset V \\ w + w' = v \longmapsto w' \end{cases}$$

όπου $v = w + w'$ είναι η μοναδική γραφή του $v \in V$ ως άθροισμα δύο στοιχείων $w \in W, w' \in W^\perp$.

Είναι σαφές ότι $\text{Id}_V = p_W + p_{W^\perp}$, $p_W^2 = p_W$, $p_{W^\perp}^2 = p_{W^\perp}$ και $p_W p_{W^\perp} = p_{W^\perp} p_W = 0$. Τέλος για τους πίνακες των γραμμικών συναρτήσεων $p_W, p_{W^\perp}$ ως προς την βάση B_V έχουμε

$$(p_W, B_V, B_V) = \left(\begin{array}{c|c} \mathbb{I}_n & \mathbf{0}_{n,m} \\ \hline \mathbf{0}_{m,n} & \mathbf{0}_{m,m} \end{array} \right) \quad (p_{W^\perp}, B_V, B_V) = \left(\begin{array}{c|c} \mathbf{0}_{n,n} & \mathbf{0}_{n,m} \\ \hline \mathbf{0}_{m,n} & \mathbb{I}_n \end{array} \right)$$

Παρατήρηση VII.6.8. Για ένα πίνακα $A = (a_{ij}) \in \mathbb{R}^{n,m}$ θεωρούμε το γραμμικό σύστημα

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,m} \\ a_{21} & a_{22} & \cdots & a_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Έστω $a_1, \dots, a_n$ οι γραμμές του A και έστω $W \subset \mathbb{R}^m$ ο χώρος γραμμών του πίνακα. Είναι σαφές ότι ο χώρος λύσεων του συστήματος ταυτίζεται με τον

$$W^\perp = \{x = (x_1, \dots, x_m)^t \in \mathbb{R}^m : \langle w, x \rangle = 0 \text{ για κάθε } w \in W\}.$$

οπότε ισχύει ότι $\dim W + \dim W^\perp = m$ και επίσης $\dim \operatorname{Im} A + \dim W^\perp = m$, και αφού $\dim \operatorname{Im}(A)$ είναι η διάσταση του χώρου στηλών, έχουμε μια διαφορετική απόδειξη ότι η διάσταση του χώρου γραμμών ταυτίζεται με αυτή του χώρου στηλών.

Ασκήσεις

Άσκηση VII.6 Δίνεται ένα σύνολο $u_1, \dots, u_n$ ορθογώνιων διανυσμάτων σε διανυσματικό χώρο V με εσωτερικό γινόμενο. Για κάθε διάνυσμα $v \in V$ ισχύει

$$\sum_{v=1}^n \frac{|\langle v, u_v \rangle|^2}{\|u_v\|^2} \leq \|v\|^2$$

με την ισότητα να ισχύει μόνο αν

$$v = \sum_{v=1}^n \frac{\langle v, u_v \rangle}{\|u_v\|^2} u_v.$$

Άσκηση VII.7 Με το εσωτερικό γινόμενο στον χώρο των συνεχών συναρτήσεων $C[0, 1] = \{f : [0, 1] \rightarrow \mathbb{R}, f \text{ συνεχής}\}$ να βρεθεί μια ορθοκανονική βάση του χώρου που παράγουν τα $\{1, x, x^2, x^3\}$.

Άσκηση VII.8 Θεωρούμε τον διανυσματικό χώρο $\mathbb{C}^{n,n}$, με το εσωτερικό γινόμενο $\langle A, B \rangle = \operatorname{tr}(AB^*)$, όπου για πίνακα $B = (b_{ij})$ ο πίνακας $B^* = \overline{B}^t = (\overline{b}_{ji})$. Να βρεθεί το ορθογώνιο συμπλήρωμα του χώρου των διαγωνίων πινάκων.

Άσκηση VII.9 Δίνεται ο πίνακας

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3 \times 3},$$

του οποίου το χαρακτηριστικό πολυώνυμο δίνεται ότι είναι το $\chi_A = -(1-x)x(3-x)$.

1. Να βρεθεί μια βάση για κάθε ιδιόχωρο του A .
2. Να βρεθεί μια ορθοκανονική βάση του $\mathbb{R}^{3 \times 1}$, αποτελούμενη από ιδιοδιανύσματα του A .
3. Να βρεθεί πίνακας $\Gamma \in \mathbb{R}^{3 \times 3} \neq 0$ ώστε $\Gamma A = A \Gamma = 0$.

VII.6.3 Η μέθοδος των ελαχίστων τετραγώνων

Υποθέτουμε ότι μια ποσότητα εξαρτάται με γραμμικό τρόπο από τις μεταβλητές $x_1, \dots, x_n$ δηλαδή εκφράζεται ως συνάρτηση

$$f(x_1, x_2, \dots, x_n) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n.$$

Θέλουμε να υπολογίσουμε τα $a_1, \dots, a_n$. Σε ένα κόσμο που μπορούμε να κάνουμε μετρήσεις με όσο μεγάλη ακρίβεια θέλουμε θα αρκούσε να εκτελέσουμε n το πλήθος πειράματα μετρώντας

κάθε φορά τις τιμές των $x_1, \dots, x_n, f(x_1, x_2, \dots, x_n)$ και στην συνέχεια να λύσουμε το σύστημα με αγνώστους $a_1, \dots, a_n$. Το πρόβλημα σε αυτή την προσέγγιση είναι ότι δεν είναι δυνατόν ούτε να μετρήσουμε με απόλυτη ακρίβεια τα $x_1, \dots, x_n, f(x_1, \dots, x_n)$. Για αυτό τον λόγο κάνουμε μια σειρά πειραμάτων, πολύ περισσότερα από n και στην συνέχεια προσπαθούμε να προσδιορίσουμε όσο γίνεται καλύτερα τις τιμές $a_1, \dots, a_n$ με τον ακόλουθο τρόπο.

Υποθέτουμε ότι στο i πείραμα έχουμε τις τιμές $x_{1i}, \dots, x_{ni}$ και $y_i = f(x_{1i}, \dots, x_{ni})$ για $1 \leq i \leq m$, $m > n$ οπότε έχουμε το σύστημα

$$\begin{pmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Το παραπάνω σύστημα είναι αδύνατο και αυτό οφείλεται στα λάθη τα οποία υπεισέρχονται στις πειραματικές μετρήσεις. Θέλουμε να προσδιορίσουμε τα $a_1, \dots, a_n$ ελαχιστοποιώντας την ποσότητα

$$\sum_{i=1}^m (a_1 x_{1i} + a_2 x_{2i} + \cdots + a_n x_{ni} - y_i)^2.$$

Αν τον πίνακα

$$X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{pmatrix}$$

τότε ζητάμε να διαλέξουμε διάνυσμα $a = (a_1, \dots, a_n)^t$ ώστε να ελαχιστοποιείται η απόσταση $\|Xa - y\|$, όπου $y = (y_1, \dots, y_m)^t$. Δηλαδή ψάχνουμε να βρούμε την ελάχιστη απόσταση του y από τον χώρο που παράγουν οι στήλες του πίνακα. Με άλλα λόγια το σύστημα είναι αδύνατο δηλαδή το y δεν ανήκει στην εικόνα της γραμμικής συνάρτησης που ορίζει ο πίνακας X , οπότε ψάχνουμε να βρούμε το πλησιέστερο διάνυσμα της εικόνας στο y . Συνεπώς αν $x^1, \dots, x^{n'}$, $n \leq n'$ είναι μια βάση του χώρου στηλών του πίνακα X ψάχνουμε για $v = X'a$ στον χώρο στηλών ώστε

$$\langle v - y, x^i \rangle = 0 \text{ για κάθε } i = 1, \dots, n'.$$

Δηλαδή έχουμε το σύστημα

$$a_1 \langle x^1, x^i \rangle + a_2 \langle x^2, x^i \rangle + \cdots + a_{n'} \langle x^{n'}, x^i \rangle = \langle y, x^i \rangle \text{ για } i = 1, \dots, n', \quad (\text{VII.17})$$

με

$$\langle y, x^i \rangle = \sum_{v=1}^m x_{iv} y_v \quad \langle x^j, x^i \rangle = \sum_{v=1}^m x_{jv} x_{iv}.$$

αυτό είναι ένα σύστημα n' -εξισώσεων με n' -αγνώστους το οποίο έχει μοναδική λύση αρκεί ο πίνακας

$$A = (a_{ij}) \text{ με } a_{ij} = \langle x^j, x^i \rangle$$

να είναι αντιστρέψιμος. Παρατηρούμε ότι ο πίνακας $A = (X')^t X'$, όπου X' είναι ο πίνακας που έχει σαν στήλες τα $x^1, \dots, x^{n'}$. Στο επόμενο λήμμα θα δείξουμε ότι ο πίνακας αυτός είναι αντιστρέψιμος και συνεπώς έχουμε ότι οι λύσεις το συστήματος (VII.17) δίνονται από

$$(a_1, \dots, a_{n'})^t = ((X')^t X')^{-1} (X')^t (y_1, \dots, y_m)^t.$$

Λήμμα VII.6.9. Έστω πίνακας X' πίνακας $m \times n'$ με στήλες τις $x^1, \dots, x^{n'}$. Το σύνολο στηλών $x^1, \dots, x^{n'}$ είναι γραμμικά ανεξάρτητα αν και μόνο αν ο πίνακας $(X')^t(X')$ είναι αντιστρέψιμος.

Απόδειξη. Ας υποθέσουμε ότι ο $n' \times n'$ πίνακας $(X')^t X'$ δεν είναι αντιστρέψιμος. Τότε υπάρχει μη-μηδενικό $v \in \mathbb{R}^{n',1}$ με $(X')^t X'v = 0_{n',1}$ και συνεπώς και το $v^t (X')^t X'v = 0$ άρα $\langle X'v, X'v \rangle = \|X'v\|^2 = 0$ το οποίο μας δίνει ότι $X'v = 0_{\mathbb{R}^m}$ συνεπώς και ο πίνακας X' έχει μη τετριμμένο πυρήνα, άρα οι στήλες του είναι γραμμικά εξαρτημένες.

Αντιστρόφως, κάθε σχέση εξάρτησης μεταξύ των στηλών δίνει ένα μη-μηδενικό $w \in \mathbb{R}^{n'}$, ώστε $'w = 0$. Άρα $(X')^t Xw = 0$ συνεπώς ο $n' \times n'$ πίνακας $(X')^t X$ δεν μπορεί να είναι αντιστρέψιμος. $\square$

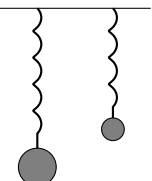
Πόρισμα VII.6.10. Στην περίπτωση που έχουμε μια εξίσωση της μορφής $f(x) = ax$ και ζητούμε να προσδιορίσουμε το a , δηλαδή θέλουμε να βρούμε την «καλύτερη» ευθεία της μορφής $y = ax$ που να προσεγγίζει τα σημεία $(x_1, y_1), \dots, (x_n, y_n)$ τότε θέτουμε $x = (x_1, \dots, x_n)^t$ $y = (y_1, \dots, y_n)^t$ και το a δίνεται από το

$$a = (x^t x)^{-1} (x^t y) = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

Παράδειγμα VII.6.11. Είναι γνωστό ότι η δύναμη που ασκεί ένα ελατήριο όταν τεντωθεί είναι ανάλογη της μετατόπισης, δηλαδή

$$F = -kx$$

για κάποια σταθερά k . Θέλουμε να την μετρήσουμε την σταθερά k για κάποιο συγκεκριμένο ελατήριο. Υποβάλλουμε το ελατήριο σε μια σειρά από παραμορφώσεις και μετράμε την δύναμη που ασκεί αυτό και πειραματικά λογαριάζουμε τον πίνακα: Ας θεωρήσουμε ότι έχουμε τις εξής μετρήσεις:

	Αποκλίση (x)	Δύναμη (F)
	0.0	0.0
	0.1	5.2
	0.2	10.5
	0.3	15.7
	0.4	20.9

από όπου υπολογίζουμε ότι

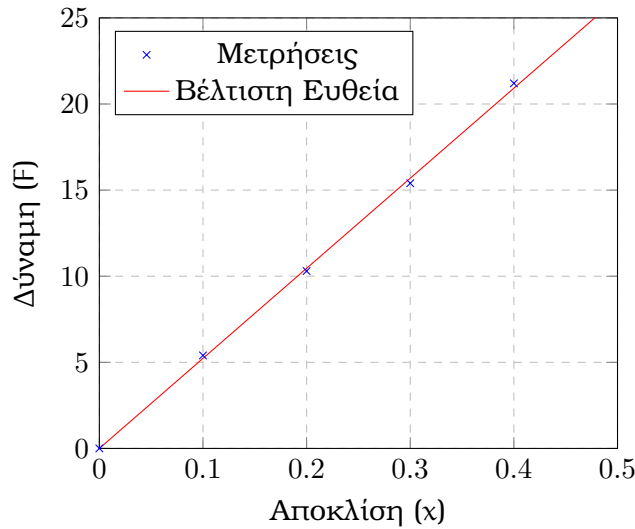
$$\sum_{i=1}^5 x_i^2 = 0.01 + 0.04 + 0.09 + 0.16 = 0.3$$

$$\sum_{i=1}^5 F_i x_i = 0.0 + 0.52 + 2.1 + 4.71 + 8.36 = 15.69$$

από όπου έχουμε ότι $k = \frac{15.69}{0.3} \approx 52.3$

Παράδειγμα VII.6.12. Η πλησιέστερη έλλειψη σε μία σειρά σημείων. Αν μας δοθούν μια σειρά σημείων στον επίπεδο μπορούμε να υπολογίσουμε την πλησιέστερη έλλειψη σε αυτά;

Αυτό είναι ένα πρόβλημα το οποίο έλυσε ο Gauss σε ηλικία 24 ετών με την μέθοδο των ελαχίστων τετραγώνων την οποία είχε επινοήσει ο ίδιος. Οι αστρονόμοι είχαν παρατηρήσει ένα νέο πλανήτη τον αστεροειδή «Δήμητρα» και είχαν δεδομένα για την θέση του μια χρονική στιγμή



Σχήμα VII.7: Μετρήσεις και βέλτιστη ευθεία

αλλά δεν ήταν δυνατόν να τον ξανά παρατηρήσουν. Ο Gauss υπολόγισε την τροχιά της Δήμητρας και είπε στους αστρονόμους πότε και που να στρέψουν τα τηλεσκόπια τους ώστε να ξαναδούν τον χαμένο αστεροειδή. Η ανακάλυψη της Δήμητρας ήταν ένας θρίαμβος των Μαθηματικών και του νεαρού τότε Gauss. Ο Gauss έλυσε το πρόβλημα στον χώρο, εμείς για χάρη ευκολίας θα περιοριστούμε σε ένα πρόβλημα στο επίπεδο.

Μια έλλειψη δίνεται ως ένα σύνολο σημείων $(x, y) \in \mathbb{R}^2$ που ικανοποιούν μια εξίσωση της μορφής

$$x^2 + By^2 + Cxy + Dx + Ey + F = 0.$$

Η παραπάνω εξίσωση μπορεί να αντιστοιχεί και σε άλλες κωνικές τομές, κύκλους, παραβολές και υπερβολές. Αν μας δοθούν τα σημεία $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ τότε υπάρχει μια έλλειψη η οποία περνάει από αυτά αν και μόνο αν

$$x_i^2 + By_i^2 + Cx_iy_i + Dx_i + Ey_i + F = 0 \text{ για κάθε } 1 \leq i \leq n.$$

Μπορεί να υπάρχουν λάθη στις συντεταγμένες των σημείων και να μην ανήκουν όλα σε μια εξίσωση της παραπάνω μορφής. Ο στόχος μας είναι να υπολογίσουμε την πλησιέστερη κωνική τομή στα δεδομένα σημεία. Το πρόβλημα μπορεί να γραφεί σε μορφή πινάκων ως εξής

$$\begin{pmatrix} y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_n^2 & x_ny_n & x_n & y_n & 1 \end{pmatrix} \begin{pmatrix} B \\ C \\ D \\ E \\ F \end{pmatrix} = - \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix}$$

Στο παραπάνω πρόβλημα τα x_i, y_i είναι γνωστά και ζητούμε να υπολογίσουμε τα B, C, D, E, F . Έστω

$$A = \begin{pmatrix} y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_n^2 & x_ny_n & x_n & y_n & 1 \end{pmatrix} \text{ και } b = - \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix}$$

Έστω το σύνολο των σημείων του επιπέδου

$$\{(8, 2), (-6.07, -3.18), (-1.53, -6.82), (4.39, -4.74), (5.27, 0.66), (3, 5), (-5.31, 6.38)\}.$$

Υπολογίζουμε τους πίνακες A , b

$$A = \begin{pmatrix} 4 & -16 & -8 & 2 & 1 \\ 10.1124 & 19.3026 & -6.07 & -3.18 & 1 \\ 46.5124 & 10.4346 & -1.53 & -6.82 & 1 \\ 22.4676 & -20.8086 & 4.39 & -4.74 & 1 \\ 0.4356 & 3.4782 & 5.27 & 0.66 & 1 \\ 25 & 15 & 3 & 5 & 1 \\ 40.7044 & -33.8778 & -5.31 & 6.38 & 1 \end{pmatrix} \text{ και } b = \begin{pmatrix} -64 \\ -36.8449 \\ -2.3409 \\ -19.2721 \\ -27.7729 \\ -9 \\ -28.1961 \end{pmatrix}$$

Στην συνέχεια υπολογίζουμε

$$A^t A = \begin{pmatrix} 5068.49 & -853.446 & -204.758 & -62.8869 & 149.232 \\ -853.446 & 2555.27 & 146.74 & -204.758 & -22.471 \\ -204.758 & 146.74 & 187.427 & -22.471 & -8.25 \\ -62.8869 & -204.758 & -22.471 & 149.232 & -0.7 \\ 149.232 & -22.471 & -8.25 & -0.7 & 7 \end{pmatrix}$$

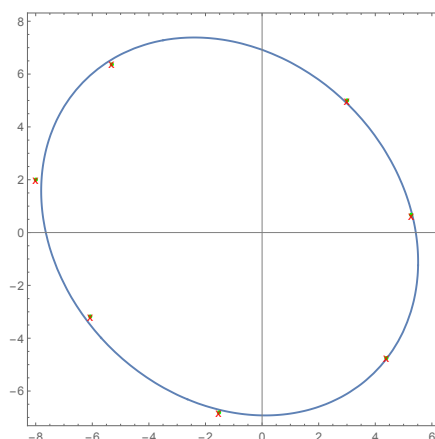
Τέλος έχουμε ότι

$$\begin{pmatrix} B \\ C \\ D \\ E \\ F \end{pmatrix} = (A^t A)^{-1} A^t b = \begin{pmatrix} 0.866743 \\ 0.350403 \\ 2.21155 \\ 0.000974531 \\ -41.5219 \end{pmatrix}$$

όποτε καταλήγουμε στην εξίσωση της έλλειψης:

$$x^2 + 0.866743y^2 + 0.350403xy + 2.21155x + 0.000974531y - 41.5219 = 0.$$

Στο σχήμα **VII.8** έχουμε σχεδιάσει την παραπάνω εξίσωση μαζί με τα αρχικά σημεία.



Σχήμα VII.8: Η βέλτιστη έλλειψη από μία επιλογή σημείων

VII.7 Η συζυγής γραμμική συνάρτηση

Ορισμός VII.7.1. Ο δυικός χώρος ενός $\mathbb{F}$ -διανυσματικού χώρου V είναι ο χώρος των γραμμικών συναρτήσεων

$$\text{Hom}(V, \mathbb{F}) := \{f : V \rightarrow \mathbb{F}, f \text{ είναι γραμμική}\}.$$

Τα στοιχεία του δυϊκού χώρου λέγονται και γραμμικά συναρτησοειδή.

Είναι σαφές ότι είναι και αυτός ένας διανυσματικός χώρος με πράξεις

$$\begin{aligned} + : V^* \times V^* &\longrightarrow V^* \\ (f, g) &\longmapsto f + g : v \rightarrow (f + g)(v) = f(v) + g(v) \end{aligned}$$

$$\begin{aligned} \cdot : \mathbb{F} \times V^* &\longrightarrow V^* \\ (\lambda, g) &\longmapsto \lambda g : v \rightarrow (\lambda f)(v) = \lambda f(v) \end{aligned}$$

Παρατήρηση VII.7.2. Κάθε στοιχείο $a \in V$ ορίζει ένα στοιχείο στο V^* , το $v \mapsto \langle v, a \rangle$. Θα δείξουμε ότι σε πεπερασμένης διάστασης χώρους ισχύει και το αντίθετο.

Πρόταση VII.7.3. Θεωρούμε ένα διανυσματικό χώρο πεπερασμένης διάστασης V και $f \in V^*$. Τότε υπάρχει μοναδικό $a \in V$ ώστε

$$f(v) = \langle v, a \rangle, \text{ για κάθε } v \in V.$$

Απόδειξη. Θεωρούμε μια ορθοκανονική βάση $\{e_1, \dots, e_n\}$ του V . Θέτουμε

$$a = \sum_{j=1}^n \overline{f(e_j)} e_j,$$

και υπολογίζουμε

$$\langle e_i, a \rangle = \left\langle e_i, \sum_{j=1}^n \overline{f(e_j)} e_j \right\rangle = f(e_i)$$

από όπου προκύπτει το ζητούμενο. □

Παρατήρηση VII.7.4. Σε άπειρης διάστασης διανυσματικούς χώρους η παραπάνω πρόταση δεν είναι αληθής. Θεωρούμε τον διανυσματικό χώρο $V = \mathbb{C}[x]$, με εσωτερικό γινόμενο

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

Όπως είδαμε στην άσκηση VII.5 για $f = \sum_{\nu} a_{\nu} x^{\nu}$, $g = \sum_{\mu} b_{\mu} x^{\mu}$

$$\langle f, g \rangle = \sum_{\nu, \mu} a_{\nu} \overline{b_{\mu}} \frac{1}{\nu + \mu + 1}.$$

Για ένα σταθερό $z \in \mathbb{C}$ Θεωρούμε το στοιχείο του V^*

$$\begin{aligned} \mathbb{C}[x] &\longrightarrow \mathbb{C} \\ f &\longmapsto \text{eval}(f) = f(z). \end{aligned}$$

Αυτή είναι μια γραμμική συνάρτηση δηλαδή

$$\text{eval}(f + g) = (f + g)(z) = f(z) + g(z) = \text{eval}(f) + \text{eval}(g) \text{ για κάθε } f, g \in \mathbb{C}[x]$$

και

$$\text{eval}(\lambda f) = \lambda f(z) = \lambda \text{eval}(f) \text{ για κάθε } \lambda \in \mathbb{C}, f \in \mathbb{C}[x].$$

Υπάρχει πολυώνυμο $g \in \mathbb{C}[x]$ ώστε

$$\text{eval}(f) = \langle f, g \rangle \text{ για κάθε } f \in \mathbb{C}[x];$$

Η απάντηση είναι όχι. Πράγματι, υποθέτουμε ότι το $f(x) = (x - z)f_1(x)$. Τότε έχουμε

$$0 = f(z) = \text{eval}(f) = \int_0^1 f(t)\overline{g(t)}dt = \int_0^1 (x - z)f_1(t)\overline{g(t)}dt,$$

και η τελευταία σχέση είναι αληθής για κάθε $f_1(x)$ άρα και για $f_1(x) = \overline{x - zg(x)}$, το οποίο δίνει ότι

$$0 = \int_0^1 |t - z|^2 |g(t)|^2 dt$$

το οποίο μπορεί να συμβεί μόνο αν $g = 0$. Όμως το eval δεν είναι το μηδενικό συναρτησοειδές, άτοπο.

Το αποτέλεσμα είναι αναμενόμενο, δεν είναι δυνατόν ένα πολυώνυμο g , το οποίο έχει πεπερασμένη πληροφορία διάστασης $\deg g + 1$ να περιγράψει μια γραμμική συνάρτηση που έχει πληροφορία στο πως δρα σε ένα άπειρης διάστασης χώρο. Στο μάθημα της συναρτησιακής ανάλυσης θα δείτε ότι με επιπλέον συνθήκες, για παράδειγμα την πληρότητα της μετρικής, δηλαδή εξασφαλίζοντας ότι κάθε ακολουθία Cauchy ως προς την μετρική $\|v\| = \langle v, v \rangle^{1/2}$ συγκλίνει, έχουμε ένα θεώρημα αναπαράστασης των “συνεχών” συναρτησοειδών f στην μορφή $f(\cdot) = \langle \cdot, v \rangle$ και για απειροδιάστατους χώρους¹.

Παρατήρηση VII.7.5. Έστω $f : V \rightarrow \mathbb{F}$ ένα στοιχείο του δυικού χώρου. Θέτουμε $W = \ker(f)$. Αν το $f \neq 0$, τότε $\text{rk}(f) = 1$ και $\dim W = \dim V - 1$. Θεωρούμε την ανάλυση $V = W \oplus W^\perp$. Η τιμή του f προσδιορίζεται από την τιμή στο $W^\perp$. Αν $p_{W^\perp} : V \rightarrow W^\perp$ είναι η ορθογώνια προβολή, τότε $f(v) = f(p_{W^\perp}(v))$ για κάθε $v \in V$. Αν $f \neq 0$, και w ένα μη μηδενικό διάνυσμα του $W^\perp$, τότε

$$p_{W^\perp}(v) = \frac{\langle v, w \rangle}{\|w\|^2} w \text{ για κάθε } v \in V.$$

Συνεπώς,

$$f(v) = \langle v, w \rangle \frac{f(w)}{\|w\|^2} = \left\langle v, \frac{\overline{f(w)}}{\|w\|^2} w \right\rangle \text{ για κάθε } v \in V.$$

Λήμμα VII.7.6. Αν για κάθε $v \in V$ ισχύει

$$\langle v, w_1 \rangle = \langle v, w_2 \rangle$$

τότε $w_1 = w_2$.

Απόδειξη. Έχουμε ότι για κάθε $v \in V$ $\langle v, w_1 - w_2 \rangle = 0$, οπότε αρκεί να πάρουμε $v = w_1 - w_2$. $\square$

Θεώρημα VII.7.7. Έστω V πεπερασμένης διάστασης διανυσματικός χώρος με εσωτερικό γινόμενο. Για κάθε γραμμική συνάρτηση $T : V \rightarrow V$, υπάρχει μοναδική γραμμική συνάρτηση T^* ώστε

$$\langle Tv, w \rangle = \langle v, T^*w \rangle \text{ για κάθε } v, w \in V.$$

Απόδειξη. Σταθεροποιούμε ένα $w \in V$. Τότε η συνάρτηση

$$\begin{aligned} V &\longrightarrow \mathbb{F} \\ v &\longmapsto \langle Tv, w \rangle \end{aligned}$$

¹https://en.wikipedia.org/wiki/Riesz_representation_theorem

είναι ένα στοιχείο του δυικού χώρου και συνεπώς υπάρχει μοναδικό $w' \in V$, το οποίο θα το συμβολίζουμε με T^*w , ώστε

$$\langle Tv, w \rangle = \langle v, T^*w \rangle.$$

Θα πρέπει να δείξουμε ότι το T^*w είναι μια γραμμική συνάρτηση του w . Πράγματι, για $\lambda_1, \lambda_2 \in \mathbb{F}$ και $w_1, w_2 \in V$ έχουμε

$$\begin{aligned} \langle v, T^*(\lambda_1 w_1 + \lambda_2 w_2) \rangle &= \langle Tv, \lambda_1 w_1 + \lambda_2 w_2 \rangle \\ &= \overline{\lambda_1} \langle Tv, w_1 \rangle + \overline{\lambda_2} \langle Tv, w_2 \rangle \\ &= \overline{\lambda_1} \langle v, T^*w_1 \rangle + \overline{\lambda_2} \langle v, T^*w_2 \rangle \\ &= \langle v, \lambda_1 T^*(w_1) + \lambda_2 T^*(w_2) \rangle, \end{aligned}$$

από όπου έχουμε, σύμφωνα με το λήμμα VII.7.6, ότι

$$T^*(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 T^*(w_1) + \lambda_2 T^*(w_2).$$

□

Θεώρημα VII.7.8. Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V θεωρούμε τον χώρο

$$\text{End}(V) = \{T : V \rightarrow V : T \text{ γραμμική}\}.$$

Η συνάρτηση $T \mapsto T^*$ είναι μια συνάρτηση $\text{End}(V) \rightarrow \text{End}(V)$ η οποία ικανοποιεί

1. $(T_1 + T_2)^* = T_1^* + T_2^*$, για κάθε $T_1, T_2 \in \text{End}(V)$
2. $(\lambda T)^* = \overline{\lambda} T^*$, για κάθε $\lambda \in \mathbb{C}$ και κάθε $T \in \text{End}(V)$
3. $(T_1 T_2)^* = T_2^* T_1^*$ για κάθε $T_1, T_2 \in \text{End}(V)$
4. $(T^*)^* = T$ για κάθε $T \in \text{End}(V)$

Απόδειξη. Για το 1. παρατηρούμε ότι για κάθε $v, w \in V$ έχουμε

$$\langle v, (T_1 + T_2)^* w \rangle = \langle (T_1 + T_2)v, w \rangle = \langle T_1 v, w \rangle + \langle T_2 v, w \rangle = \langle v, T_1^* w \rangle + \langle v, T_2^* w \rangle = \langle v, (T_1^* + T_2^*) w \rangle$$

οπότε η ισότητα προκύπτει λόγω του μονοσημάντου του $(T_1 + T_2)^*$.

Το 2. προκύπτει με παρόμοιο τρόπο

$$\langle v, (\lambda T)^* w \rangle = \langle \lambda T v, w \rangle = \lambda \langle T v, w \rangle = \lambda \langle v, T^* w \rangle = \langle v, \overline{\lambda} T^* w \rangle.$$

Για το 3. έχουμε

$$\langle v, (T_1 T_2)^* w \rangle = \langle T_1 T_2 v, w \rangle = \langle T_2 v, T_1^* w \rangle = \langle v, T_2^* T_1^* w \rangle.$$

Για το 4. υπολογίζουμε

$$\langle v, (T^*)^* w \rangle = \langle T^* v, w \rangle = \overline{\langle w, T^* v \rangle} = \overline{\langle Tw, v \rangle} = \langle v, Tw \rangle.$$

□

Πρόταση VII.7.9. Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο με εσωτερικό γινόμενο και $B = \{e_1, \dots, e_n\}$ μια διατεταγμένη ορθοκανονική βάση, αν $(T, B, B) = (a_{ij})$ είναι ο πίνακας μιας γραμμικής συνάρτησης $T : V \rightarrow V$, έχουμε ότι $a_{ij} = \langle Te_j, e_i \rangle$.

Απόδειξη. Εκφράζουμε κάθε $v \in V$ ως γραμμικό συνδυασμό των στοιχείων της ορθοκανονικής βάσης:

$$v = \sum_{i=1}^n \langle v, e_i \rangle e_i.$$

Ο πίνακας του T υπολογίζεται από την σχέση

$$Te_j = \sum_{i=1}^n a_{i,j} e_i$$

άρα

$$\langle Te_j, e_i \rangle = \sum_{i=1}^n \langle a_{i,j} e_i, e_i \rangle = a_{ij}.$$

□

Πόρισμα VII.7.10. Σε κάθε ορθοκανονική βάση B του πεπερασμένης διάστασης χώρου V , και κάθε γραμμική $T : V \rightarrow V$ ισχύει ότι

$$(T^*, B, B) = (T, B, B)^*,$$

όπου για ένα πίνακα A θέτουμε $A^* = \overline{A}^t$, δηλαδή τον ανάστροφο στον οποίο έχουμε εφαρμόσει την μιγαδική συζυγία σε κάθε στοιχείο του.

Απόδειξη. Έστω $B = \{e_1, \dots, e_n\}$ μία ορθοκανονική βάση του V και $A = (a_{ij}) = (T, B, B)$, $B = (b_{ij}) = (T^*, B, B)$. Έχουμε ότι

$$a_{ij} = \langle Te_j, e_i \rangle, \quad b_{ij} = \langle T^* e_j, e_i \rangle.$$

Συνεπώς

$$b_{ij} = \langle T^* e_j, e_i \rangle = \langle e_j, Te_i \rangle = \overline{\langle Te_i, e_j \rangle} = \overline{a_{ji}},$$

από όπου προκύπτει το ζητούμενο. □

Παράδειγμα VII.7.11. Αν θεωρήσουμε τον χώρο $\mathbb{C}^n$ των διανυσμάτων στηλών, τότε η συνάρτηση

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (x, y) &\longmapsto y^* x = \langle x, y \rangle \end{aligned}$$

ορίζει το ερμητιανό εσωτερικό γινόμενο. Μια γραμμική συνάρτηση δίνεται μέσω ενός $n \times n$ πίνακα A , δηλαδή $x \mapsto Ax$. Έχουμε ότι

$$\langle Ax, y \rangle = y^* Ax = (A^* y)^* x = \langle x, A^* y \rangle.$$

Ασκήσεις

Άσκηση VII.10 Θεωρούμε τον χώρο $\mathbb{C}^{n,n}$ με το εσωτερικό γινόμενο

$$\langle A, B \rangle = \text{tr}(B^* A).$$

Για ένα σταθερό πίνακα M θεωρούμε την γραμμική συνάρτηση

$$\begin{aligned} L_M : \mathbb{C}^{n,n} &\longrightarrow \mathbb{C}^{n,n} \\ A &\longmapsto MA \end{aligned}$$

Δείξτε ότι $L_M^* = L_{M^*}$.

Άσκηση VII.11 Θεωρούμε τον διανυσματικό χώρο $\mathbb{C}[x]$ με το εσωτερικό γινόμενο

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

- Δείξτε ότι για σταθερό $f \in \mathbb{C}[x]$ για την γραμμική συνάρτηση

$$\begin{aligned} M_f : \mathbb{C}[x] &\longrightarrow \mathbb{C}[x] \\ g &\longmapsto fg \end{aligned}$$

ισχύει ότι $M_f^* = M_{\overline{f}}$.

- Για την γραμμική συνάρτηση

$$\begin{aligned} D : \mathbb{C}[x] &\longrightarrow \mathbb{C}[x] \\ g &\longmapsto Dg = g' \end{aligned}$$

δεν υπάρχει γραμμική συνάρτηση D^* , ώστε $\langle Df, g \rangle = \langle f, D^*g \rangle$ για κάθε $f, g \in \mathbb{C}[x]$.

VII.8 Μοναδιαίες γραμμικές συναρτήσεις

Ορισμός VII.8.1. Θα λέμε ότι μία γραμμική συνάρτηση $T : V \rightarrow W$ ανάμεσα σε διανυσματικούς χώρους V, W διατηρεί το εσωτερικό γινόμενο όταν

$$\langle Tv, Tv \rangle = \langle v, v \rangle, \text{ για κάθε } v \in V.$$

Ένας ισομορφισμός ανάμεσα σε χώρους με εσωτερικό γινόμενο θα είναι ένας ισομορφισμός που επιπλέον διατηρεί το εσωτερικό γινόμενο.

Παρατήρηση VII.8.2. Μια γραμμική συνάρτηση $V \rightarrow W$ που διατηρεί το εσωτερικό γινόμενο δεν μπορεί να έχει μη τετριμμένο πυρήνα, αφού $\|Tv\| = \|v\|$ και αν $v \in \ker$ τότε $v = 0$.

Συνεπώς αν $\dim V = \dim W$ κάθε γραμμική συνάρτηση $V \rightarrow W$ που διατηρεί το εσωτερικό γινόμενο είναι ισομορφισμός.

Παρατήρηση VII.8.3. Κάθε γραμμική συνάρτηση $T : V \rightarrow W$ που διατηρεί το εσωτερικό γινόμενο μεταφέρει ορθοκανονικές βάσεις σε ορθοκανονικές βάσεις. Αντιστρόφως αν $\{e_1, \dots, e_n\}$ είναι μια ορθοκανονική βάση του V , ώστε $\{Te_1, \dots, Te_n\}$ να είναι μια ορθοκανονική βάση του W , τότε για κάθε $v, w \in V$, έχουμε

$$v = \sum_{i=1}^n \lambda_i e_i, w = \sum_{j=1}^n \mu_j e_j$$

και

$$\langle v, w \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \overline{\mu_j} \langle e_i, e_j \rangle = \sum_{i=1}^n \lambda_i \overline{\mu_i}.$$

Ομοίως

$$\langle Tv, Tw \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \overline{\mu_j} \langle Te_i, Te_j \rangle = \sum_{i=1}^n \lambda_i \overline{\mu_i},$$

δηλαδή

$$\langle v, w \rangle = \langle Tv, Tw \rangle.$$

Ορισμός VII.8.4. Μία μοναδιαία (unitary) γραμμική συνάρτηση είναι ένας ισομορφισμός που διατηρεί το εσωτερικό γινόμενο.

Παρατήρηση VII.8.5. Η σύνθεση δύο μοναδιαίων γραμμικών συναρτήσεων είναι μοναδιαία γραμμική συνάρτηση, η ταυτοτική συνάρτηση είναι μοναδιαία γραμμική συνάρτηση. Συνεπώς οι μοναδιαίες γραμμικές συναρτήσεις αποτελούν υποομάδα της ομάδας των αντιστρέψιμων γραμμικών συναρτήσεων.

Πρόταση VII.8.6. Μία γραμμική συνάρτηση είναι μοναδιαία αν και μόνο αν είναι αντιστρέψιμη και $U^{-1} = U^*$.

Απόδειξη. Υποθέτουμε ότι U είναι μοναδιαία, δηλαδή διατηρεί το εσωτερικό γινόμενο. Παρατηρούμε ότι για κάθε $v, w \in W$ και $U : V \rightarrow U$ μοναδιαία έχουμε

$$\langle Uv, w \rangle = \langle Uv, \text{Id}_V w \rangle \stackrel{\text{Id}_V = U U^{-1}}{=} \langle Uv, U U^{-1} w \rangle = \langle v, U^{-1} w \rangle$$

από όπου προκύπτει ότι $U^{-1} = U^*$.

Αντιστρόφως, αν $U^{-1} = U^*$, τότε

$$\langle Uv, Uw \rangle = \langle v, U^* Uw \rangle = \langle v, w \rangle \text{ για κάθε } v, w \in V.$$

□

Ορισμός VII.8.7. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ θα λέγεται μοναδιαίος (unitary) αν και μόνο αν για $A^* = \overline{A}^t$ ισχύει $AA^* = \mathbb{I}_n$.

Πρόταση VII.8.8. Σε ένα διανυσματικό χώρο V πεπερασμένης διάστασης η γραμμική συνάρτηση $T : V \rightarrow V$ είναι μοναδιαία αν και μόνο αν ο πίνακας της ως προς μια ορθοκανονική βάση είναι μοναδιαίος.

Απόδειξη. Παρατηρούμε ότι αν $A = (T, B, B)$ είναι ο πίνακας της T ως προς μια ορθοκανονική βάση B τότε ο πίνακας (T^*, B, B) της T^* είναι ο $A^* = \overline{A}^t$. □

Ορισμός VII.8.9. Ένας πραγματικός ή μιγαδικός πίνακας θα λέγεται ορθογώνιος αν $AA^t = \mathbb{I}_n$.

Παρατήρηση VII.8.10. Ένας πραγματικός πίνακας είναι ορθογώνιος αν και μόνο αν είναι μοναδιαίος.

Πρόταση VII.8.11. Ένας πίνακας είναι μοναδιαίος αν και μόνο αν οι γραμμές ή οι στήλες του αποτελούν ορθοκανονική βάση του $\mathbb{C}^n$ με το ερμητιανό εσωτερικό γινόμενο $\langle x, y \rangle = y^* x$.

Απόδειξη. Γράφουμε τον πίνακα $A = (a^1, \dots, a^n)$, όπου a^i είναι η i -στήλη. Παρατηρούμε ότι ο πίνακας $A^* = (c_{ij})$ έχει

$$c_{ij} = (a^i)^* a^j = \langle a^j, a^i \rangle.$$

Η συνθήκη $*$ $= \mathbb{I}_n$ είναι ισοδύναμη με το ότι $c_{ij} = \delta_{ij} = \langle a^j, a^i \rangle$ δηλαδή οι στήλες αποτελούν μια ορθοκανονική βάση.

Για να δείξουμε ότι και οι στήλες αποτελούν μια ορθοκανονική βάση παρατηρούμε ότι

$$A^* A = \mathbb{I}_n \Leftrightarrow (A^t)(A^t)^* = \mathbb{I}_n,$$

δηλαδή ο A είναι μοναδιαίος αν και μόνο αν ο A^t είναι μοναδιαίος. □

Η μέθοδος Gram-Schmidt έχει την παρακάτω ενδιαφέρουσα συνέπεια:

Πρόταση VII.8.12. Για κάθε αντιστρέψιμο $n \times n$ πίνακα A υπάρχει μοναδικός κάτω τριγωνικός πίνακας B με θετικά στοιχεία στην διαγώνιο ώστε BA να είναι μοναδιαίος.

Απόδειξη. Οι γραμμές $a_1, \dots, a_n$ του πίνακα είναι γραμμικά ανεξάρτητες αφού ο πίνακας A είναι αντιστρέψιμος. Εκτελούμε την μέθοδο Gram-Schmidt στις γραμμές $a_1, \dots, a_n$ και καταλήγουμε σε μία ορθογώνια βάση $u_1, \dots, u_n$ του $\mathbb{C}^n$. Κατά την διαδικασία παρατηρούμε ότι σε κάθε βήμα της αναδρομικής ορθοκανονικοποίησης τα $\{u_1, \dots, u_i\}$ αποτελούν μια ορθογώνια βάση του διανυσματικού χώρου $\langle a_1, \dots, a_i \rangle$ και ότι

$$u_k = a_k - \sum_{j=1}^{k-1} \frac{\langle a_k, u_j \rangle}{\|u_j\|^2} u_j$$

Δηλαδή, υπάρχουν μοναδικοί συντελεστές γ_{kj} ώστε

$$u_k = a_k - \sum_{j=1}^{k-1} \gamma_{kj} a_j.$$

Θεωρούμε τον unitary πίνακα U με γραμμές $u_1/\|u_1\|, \dots, u_n/\|u_n\|$ και B τον πίνακα που ορίζεται από

$$b_{kj} = \begin{cases} -\frac{\gamma_{kj}}{\|u_k\|} & \text{αν } j < k \\ \frac{1}{\|u_k\|} & \text{αν } j = k \\ 0 & \text{αν } j > k \end{cases}$$

Παρατηρούμε ότι ο B είναι κάτω τριγωνικός, έχει θετικές τιμές στην διαγώνιο και επίσης

$$\frac{u_k}{\|u_k\|} = \sum_{j=1}^n b_{kj} a_j, \text{ για κάθε } 1 \leq k \leq n,$$

δηλαδή $U = BA$.

Θα αποδείξουμε τώρα ότι η γραφή είναι μοναδική. Θεωρούμε το σύνολο $T^+(n)$ των $n \times n$ κάτω τριγωνικών πινάκων με θετικά στοιχεία στην διαγώνιο και με $U(n)$ το σύνολο των μοναδιαίων $n \times n$ πινάκων.

Υποθέτουμε ότι $B_1, B_2 \in T^+(n)$, ώστε $B_1 A, B_2 A \in U(n)$. Έχουμε ότι ο αντίστροφος κάθε μοναδιαίου πίνακα είναι μοναδιαίος και ότι το γινόμενο δύο μοναδιαίων πινάκων είναι μοναδιαίος. Συνεπώς,

$$(B_A)(B_2 A)^{-1} = B_1 B_2^{-1} \in U(n).$$

Επίσης παρατηρούμε ότι ο αντίστροφος κάθε πίνακα στο $T^+(n)$ ανήκει στο $T^+(n)$ όπως και το γινόμενο δύο πινάκων στο $T^+(n)$ ανήκει στο $T^+(n)$, δείτε την άσκηση VII.12.

Αφού $B_1 B_2^{-1} \in \mathcal{U}(n)$, έχουμε ότι $(B_1 B_2^{-1})^{-1} = (B_1 B_2^{-1})^*$. Δηλαδή ο πίνακας $(B_2^{-1})^{-1}$ είναι κάτω τριγωνικός, ως αντίστροφος στοιχείου στο $T^+(n)$ αλλά και άνω τριγωνικός ως ανάστροφος κάτω τριγωνικού. Άρα είναι διαγώνιος. Και αφού είναι μοναδιαίος τα στοιχεία της διαγώνιου του έχουν μέτρο 1 και είναι και ταυτόχρονα θετικά. Άρα $B_1 B_2^{-1} = \mathbb{I}_n$. $\square$

Παράδειγμα VII.8.13. Να βρεθεί κάτω τριγωνικός πίνακας B με θετικά στοιχεία στην διαγώνιο για τον πίνακα A

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix}$$

ώστε ο πίνακας BA να είναι μοναδιαίος.

Στο παράδειγμα VII.6.4 κάναμε μια ορθοκανονικοποίηση των γραμμών του πίνακα υπολογίζοντας ότι

$$\begin{aligned} u_1 &= v_1, u_2 = v_2 - \frac{2}{7}v_1, u_3 = v_3 - \frac{1}{2}v_1 + \frac{21}{27}u_2 \\ &= v_3 - \frac{1}{2}v_1 + \frac{21}{27}(v_2 - \frac{2}{7}v_1) \\ &= v_3 + \frac{21}{27}v_2 - \frac{13}{18}v_1 \end{aligned}$$

οπότε έχουμε τον πίνακα

$$B' = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{7} & 1 & 0 \\ \frac{13}{18} & \frac{7}{9} & 1 \end{pmatrix}$$

ώστε

$$B'A = \begin{pmatrix} 1 & 2 & 3 \\ \frac{12}{7} & \frac{3}{7} & -\frac{6}{7} \\ -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \end{pmatrix}$$

Ο τελευταίος πίνακας δεν είναι μοναδιαίος γιατί οι γραμμές είναι ορθογώνιες αλλά δεν έχουν μέτρο 1. Αντικαθιστώντας τον πίνακα B' με τον πίνακα

$$B = \begin{pmatrix} \frac{1}{\sqrt{14}} & 0 & 0 \\ -\frac{2\sqrt{7}}{7\sqrt{27}} & \sqrt{7/27} & 0 \\ \frac{13}{18}\sqrt{6} & \frac{7}{9}\sqrt{6} & \sqrt{6} \end{pmatrix}$$

έχουμε την ζητούμενη ανάλυση.

Ασκήσεις

Άσκηση VII.12 Έστω $T^+(n)$ το σύνολο των $n \times n$ κάτω τριγωνικών πινάκων με θετικά στοιχεία στην διαγώνιο. Δείξτε ότι για κάθε $B_1, B_2 \in T^+(n)$ το γινόμενο $B_1 B_2 \in T^+(n)$ και ότι για κάθε $B \in T^+(n)$ ισχύει ότι $B^{-1} \in T^+(n)$.

Άσκηση VII.13 Θεωρούμε τον χώρο $V = C[0, 1]$ των συνεχών συναρτήσεων $[0, 1] \rightarrow \mathbb{R}$, με το εσωτερικό γινόμενο

$$\langle f, g \rangle_1 = \int_0^1 f(t)g(t)dt.$$

και W τον ίδιο χώρο με το εσωτερικό γινόμενο

$$\langle f, g \rangle_2 = \int_0^1 f(t)g(t)dt.$$

Αφού δείξετε ότι το $\langle \cdot, \cdot \rangle_2$ είναι εσωτερικό γινόμενο δείξτε ότι η

$$\begin{aligned} T : W &\longrightarrow V \\ f(x) &\longmapsto xf(x) \end{aligned}$$

διατηρεί το εσωτερικό γινόμενο αλλά δεν είναι ισομορφισμός.

Άσκηση VII.14 Να βρεθούν όλοι οι $n \times n$ μοναδιαίοι και ορθογώνιοι πίνακες για $n = 1, 2$.

VII.9 Κανονικές γραμμικές συναρτήσεις

Σε αυτή την παράγραφο θα απαντήσουμε στο πότε μια γραμμική συνάρτηση $L : V \rightarrow V$, όπου ο V είναι ένας πεπερασμένης διάστασης διανυσματικός χώρος έχει μια ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα. Ας υποθέσουμε ότι $B = \{e_1, \dots, e_n\}$ είναι μια τέτοια βάση. Από την μία έχουμε ότι

$$Le_j = \lambda_j e_j$$

δηλαδή ο (L, B, B) είναι ένας διαγώνιος πίνακας, συνεπώς ο $(L^*, B, B) = (L, B, B)^*$ δηλαδή είναι και αυτός διαγωνοποιήσιμος ως προς την ίδια βάση. Αυτό σημαίνει ότι $L^*L = LL^*$, δείτε και την πρόταση VI.3.9.

Ορισμός VII.9.1. Έστω $L : V \rightarrow V$ γραμμική συνάρτηση, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος. Η L θα λέγεται κανονική (normal) αν και μόνο αν $L^*L = LL^*$.

Ορισμός VII.9.2. Έστω $L : V \rightarrow V$ γραμμική συνάρτηση, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος. Η L θα λέγεται αυτοσυζυγής ή ερμητιανή αν και μόνο αν $L^* = L$.

Παρατήρηση VII.9.3. Είναι σαφές ότι μια ερμητιανή γραμμική συνάρτηση όπως και μια μοναδιαία γραμμική συνάρτηση είναι κανονική.

Πρόταση VII.9.4. Οι ιδιοτιμές μιας ερμητιανής γραμμικής συνάρτησης είναι όλες πραγματικές. Επίσης τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι κάθετα μεταξύ τους.

Απόδειξη. Έστω v ένα ιδιοδιάνυσμα της $L \rightarrow L$ το οποίο αντιστοιχεί στην ιδιοτιμή $\lambda \in \mathbb{C}$. Έχουμε ότι

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Lv, v \rangle = \langle v, L^*v \rangle = \langle v, Lv \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle$$

και αφού $\langle v, v \rangle \neq 0$ έχουμε ότι $\lambda = \bar{\lambda}$, δηλαδή $\lambda \in \mathbb{R}$.

Από την άλλη αν $Lv = \lambda v$ και $Lw = \mu w$ με $\mu \neq \lambda$ τότε

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle Lv, w \rangle = \langle v, L^*w \rangle = \langle v, \mu w \rangle = \bar{\mu} \langle v, w \rangle = \mu \langle v, w \rangle,$$

συνεπώς $\langle v, w \rangle = 0$. □

Λήμμα VII.9.5. Έστω V ένας πεπερασμένης διάστασης διανυσματικός χώρος και $L : V \rightarrow V$ γραμμική συνάρτηση. Θεωρούμε ένα L -αναλλοίωτο υπόχωρο W . Το ορθογώνιο συμπλήρωμα $W^\perp$ είναι L^* αναλλοίωτο.

Απόδειξη. Έχουμε ότι $Lw \in W$ για κάθε $w \in W$. Έστω $w' \in W^\perp$ δηλαδή $\langle w, w' \rangle = 0$ για κάθε $w \in W$. Θέλουμε να δείξουμε ότι $L^*w' \in W^\perp$ δηλαδή $\langle w, L^*w' \rangle = 0$ για κάθε $w \in W$. Όμως $\langle w, L^*w' \rangle = \langle Lw, w' \rangle = 0$ αφού $Lw \in W$. $\square$

Θεώρημα VII.9.6. Για κάθε αυτοσυζυγή γραμμική συνάρτηση $L : V \rightarrow V$, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος υπάρχει ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα.

Απόδειξη. Υποθέτουμε ότι $\dim V \geq 1$. Πάνω από το $\mathbb{C}$, έχουμε ότι ο L έχει μια ιδιοτιμή λ η οποία είναι μάλιστα πραγματική. Το αντίστοιχο ιδιοδιάνυσμα v μπορούμε να το κανονικοποιήσουμε στο $e = v/\|v\|$ ώστε να έχει μέτρο 1. Επιπλέον αν δουλεύουμε σε πραγματικούς χώρους μπορούμε να υποθέσουμε ότι το διάνυσμα v έχει πραγματικούς συντελεστές σε οποιαδήποτε βάση, διότι αν $A = (L, B, B) \in \mathbb{R}^{n,n}$ το ιδιοδιάνυσμα έχει συντεταγμένες $[v]_B \in \mathbb{R}^n$ που είναι λύση του συστήματος $(A - \lambda I_n)x = 0$, το οποίο έχει μη τετριμμένες λύσεις από την επιλογή του λ ως ρίζα του χαρακτηριστικού πολυωνύμου.

Είναι προφανές ότι αν $\dim V = 1$ υπάρχει ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα το $e = v/\|v\|$ που μόλις κατασκευάσαμε. Θα αποδείξουμε το ζητούμενο με επαγωγή. Υποθέτουμε ότι η πρόταση είναι αληθής για όλες τις αυτοσυζυγείς γραμμικές συναρτήσεις σε διανυσματικούς χώρους διάστασης $n-1$. Έστω ότι έχουμε ένα διανυσματικό χώρο διάστασης n . Θεωρούμε την ανάλυση

$$V = \langle e \rangle \oplus \langle e \rangle^\perp$$

Από το λήμμα Είναι σαφές από το λήμμα VII.9.5 ότι ο

$$L|_{\langle e \rangle^\perp} = L^*|_{\langle e \rangle^\perp} : \langle e \rangle^\perp \longrightarrow \langle e \rangle^\perp$$

περιορίζεται στον $\langle e \rangle^\perp$ σε μια αυτοσυζυγή συνάρτηση η οποία λόγω της επαγωγικής υπόθεσης έχει μια ορθοκανονική βάση $\{e_2, \dots, e_n\}$ αποτελούμενη από ιδιοδιανύσματα. Η $\{e, e_2, \dots, e_n\}$ είναι η ζητούμενη ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα για όλο τον χώρο. $\square$

Λήμμα VII.9.7. Για μία κανονική γραμμική συνάρτηση $L : V \rightarrow V$ για $v \in V$ έχουμε $\|Lv\| = \|L^*v\|$.

Απόδειξη. Έχουμε ότι

$$\|Lv\|^2 = \langle Lv, Lv \rangle = \langle v, L^*Lv \rangle = \langle v, LL^*v \rangle = \langle L^*v, L^*v \rangle = \|L^*v\|^2.$$

$\square$

Πρόταση VII.9.8. Έστω V πεπερασμένης διάστασης διανυσματικός χώρος και $L : V \rightarrow V$ κανονική γραμμική συνάρτηση. Τότε το $v \in V$ είναι ιδιοδιάνυσμα για την L με ιδιοτιμή λ αν και μόνο αν v είναι ιδιοδιάνυσμα για την L^* με ιδιοτιμή $\bar{\lambda}$.

Απόδειξη. Για μια τιμή $\lambda \in \mathbb{C}$ έχουμε ότι η συνάρτηση $T = L - \lambda \text{Id}_V$ είναι κανονική αφού $T^* = L^* - \bar{\lambda} \text{Id}_V$. Υπολογίζουμε ότι

$$\begin{aligned} TT^* &= (L - \lambda \text{Id}_V)(L^* - \bar{\lambda} \text{Id}_V) = LL^* - \bar{\lambda}L - \lambda L^* - \lambda \bar{\lambda} \text{Id}_V \\ &= L^*L - \bar{\lambda}L - \lambda L^* - \lambda \bar{\lambda} \text{Id}_V = (L^* - \bar{\lambda} \text{Id}_V)(L - \lambda \text{Id}_V) = T^*T. \end{aligned}$$

Συνεπώς

$$\|(L - \lambda \text{Id}_V)v\| = \|(L^* - \bar{\lambda} \text{Id}_V)v\|$$

από όπου προκύπτει ότι $(L - \lambda \text{Id}_V)v = 0$ αν και μόνο αν $(L^* - \bar{\lambda} \text{Id}_V)v = 0$, δηλαδή το ζητούμενο. $\square$

Ορισμός VII.9.9. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ θα λέγεται κανονικός αν και μόνο αν $AA^* = A^*A$.

Πρόταση VII.9.10. Σε έναν πεπερασμένης διάστασης διανυσματικό χώρο V , θεωρούμε μια γραμμική συνάρτηση $L : V \rightarrow V$ ώστε ο πίνακας της (L, B, B) ως προς μία ορθοκανονική βάση B να είναι άνω τριγωνικός. Η L είναι κανονική αν και μόνο αν (L, B, B) είναι διαγώνιος πίνακας.

Απόδειξη. Έστω $A = (L, B, B)$. Γνωρίζουμε ότι $A^* = (L^*, B, B)$. Αν $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ τότε $A^* = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n)$. Είναι σαφές ότι $A^*A = AA^*$ δηλαδή ότι ο A και συνεπώς και η L είναι κανονικοί.

Αντιστρόφως, έστω ότι L είναι κανονική και έστω $B = \{e_1, \dots, e_n\}$ μια ορθοκανονική βάση του V . Έχουμε ότι e_1 είναι ιδιοδιάνυσμα αφού ο A είναι άνω τριγωνικός και μάλιστα $Le_1 = a_{11}e_1$. Από την πρόταση VII.9.8 έχουμε ότι $L^*e_1 = \bar{a}_{11}e_1$. Σύμφωνα με τον ορισμό όμως της $A^* = \bar{A}^t$ αυτό σημαίνει ότι $a_{1j} = 0$ για $j > 1$. Συνεπώς $Le_2 = a_{22}e_2$, άρα και το e_2 είναι ιδιοδιάνυσμα της L . Όπως και πριν έχουμε ότι $L^*e_2 = \bar{a}_{22}e_2$, και αυτό μας δίνει όπως και πριν ότι $a_{2j} = 0$ για $j > 2$. Συνεχίζουμε με τον ίδιο τρόπο για να δείξουμε ότι ο πίνακας A είναι διαγώνιος. $\square$

Πρόταση VII.9.11 (Διάσπαση Schur). Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V και $L : V \rightarrow V$ υπάρχει μια ορθοκανονική βάση B , ώστε ο πίνακας (L, B, B) να είναι άνω τριγωνικός.

Απόδειξη. Αν $\dim V = 1$ τότε το αποτέλεσμα είναι σαφές. Θα αποδείξουμε το ζητούμενο με επαγωγή στην διάσταση του V . Υποθέτουμε ότι η πρόταση αληθεύει για κάθε γραμμική συνάρτηση από ένα διανυσματικό χώρο διάστασης $n-1$ στον εαυτό του. Θεωρούμε ένα διανυσματικό χώρο V διάστασης n . Θεωρούμε ένα ιδιοδιάνυσμα v της L^* το οποίο το κανονικοποιούμε στο $v/\|v\|$ ώστε να έχει μέτρο 1. Ο χώρος $\langle v/\|v\| \rangle$ είναι L^* -αναλλοίωτος συνεπώς ο $\langle v/\|v\| \rangle^\perp$ είναι L -αναλλοίωτος και έχει διάσταση $n-1$, συνεπώς μπορούμε να εφαρμόσουμε την επαγωγική υπόθεση και να κατασκευάσουμε μια ορθοκανονική βάση $\{e_1, \dots, e_{n-1}\}$ στην οποία ο L να είναι άνω τριγωνικός. Η ζητούμενη βάση είναι η $B = \{e_1, \dots, e_{n-1}, v/\|v\|\}$. $\square$

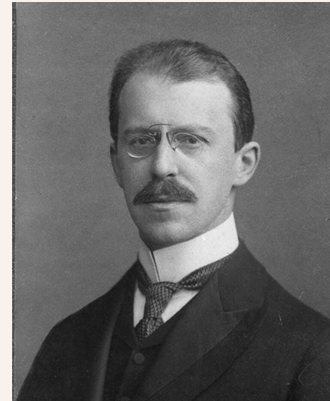
Πόρισμα VII.9.12. Για κάθε πίνακα $A \in \mathbb{C}^{n,n}$ υπάρχει μοναδιαίος πίνακας U , ώστε U^*AU να είναι άνω τριγωνικός.

Τέλος ο συνδυασμός των προτάσεων VII.9.10 και VII.9.11 δίνει το

Θεώρημα VII.9.13. Έστω ένας πεπερασμένης διάστασης διανυσματικός χώρος και $L: V \rightarrow V$ μια κανονική γραμμική συνάρτηση. Υπάρχει μια ορθοκανονική βάση του V η οποία να αποτελείται από ιδιοδιανύσματα.

Πόρισμα VII.9.14. Για κάθε κανονικό πίνακα A υπάρχει μοναδιαίος πίνακας U ώστε $U^{-1}AU$ να είναι διαγώνιος πίνακας.

Ο **Issai Schur** (10 Ιανουαρίου 1875 - 10 Ιανουαρίου 1941) ήταν Ρωσικής καταγωγής Μαθηματικός που έζησε και εργάστηκε στην Γερμανία. Ασχολήθηκε με την θεωρία ομάδων και ιδιαίτερα την θεωρία αναπαραστάσεων αλλά και με την συνδυαστική, την Θεωρία Αριθμών και την Φυσική. Ήταν μαθητής του Frobenius και στην διδακτορική διατριβή του μελέτησε τις ρητές αναπαραστάσεις της γενικής γραμμικής ομάδας. Έχει αφήσει το όνομα του σε μια σειρά από έννοιες της άλγεβρας όπως η συνθήκη του Schur, ο πολλαπλασιαστής του Schur, οι συναρτήσεις του Schur, που σχετίζονται με τις αναπαραστάσεις της συμμετρικής ομάδας και με την συνδυαστική.



Για να υπολογίσουμε ένα μοναδιαίο πίνακα U για ένα ερμητιανό πίνακα A , ώστε ο UAU^{-1} να είναι διαγώνιος παρατηρούμε ότι οι ιδιόχωροι που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι ανά δύο κάθετοι μεταξύ τους.

Μέθοδος 6 (Διαγωνιοποίηση Ερμητιανών Πινάκων).

- 1 Υπολογισμός χαρακτηριστικού πολυωνύμου
- 2 Εύρεση ιδιοτιμών $\lambda_1, \dots, \lambda_n$
- 3 Υπολογισμός βάσεων B_i των ιδιοχώρων E_{λ_i} .
- 4 Ορθοκανονικοποίηση κατά Gram-Schmidt κάθε ιδιοχώρου.

Παράδειγμα VII.9.15. Θεωρούμε τον ερμητιανό πίνακα

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

και θέλουμε να βρούμε μοναδιαίο πίνακα P , ώστε ο P^*AP να είναι διαγώνιος πίνακας.

Το χαρακτηριστικό πολυώνυμο του είναι το $\text{Ch}_A[x] = (1+x)^2(-5+x)$, το οποίο μπορούμε να το υπολογίσουμε είτε με τον ορισμό είτε ως ειδική περίπτωση της άσκησης [VI.4](#). Υπολογίζουμε ότι ο ιδιόχωρος της ιδιοτιμής $\lambda = 5$ είναι το διάνυσμα $v_1 = (1, 1, 1)^t$ ενώ ο ιδιόχωρος της ιδιοτιμής $\lambda = -1$ παράγεται από τα διανύσματα $v_2 = (1, -1, 0)^t$, $v_3 = (1, 0, -1)^t$. Θα μπορούσαμε να πάρουμε

$$Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \text{ με αντίστροφο } Q^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

και να είχαμε ότι $Q^{-1}AQ = \text{diag}(5, -1, -1)$. Όμως ο πίνακας Q δεν είναι μοναδιαίος. Για να κατασκευάσουμε τον μοναδιαίο πίνακα P ορθοκανονικοποιούμε σε $e_1 = v_1/\|v_1\| = \frac{1}{\sqrt{3}}(1, 1, 1)^t$, $e_2 = v_2/\|v_2\| = 1/\sqrt{2}(1, -1, 0)$, ενώ $e_3 = u_3/\|u_3\|$, όπου

$$u_3 = v_3 - \frac{\langle v_2, v_3 \rangle}{\langle v_2, v_2 \rangle} v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \end{pmatrix} \text{ συνεπώς } e_3 = \sqrt{2/3} \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \end{pmatrix}$$

Σχηματίζουμε τον πίνακα

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\sqrt{\frac{2}{3}} \end{pmatrix} \text{ με } P^{-1} = P^* = P^t = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\sqrt{\frac{2}{3}} \end{pmatrix}$$

και έχουμε ότι $P^*AP = \text{diag}(5, -1, -1)$.

Ασκήσεις

Άσκηση VII.15 Έστω $A \in \mathbb{C}^{n,n}$. Για ένα πολυώνυμο $f(x) = \sum_{v=0}^n a_v x^v$ θα συμβολίζουμε με $\bar{f}(x) = \sum_{v=0}^n \bar{a}_v x^v$. Αποδείξτε ότι

1. $\text{ch}_{A^*}(x) = \overline{\text{ch}_A(x)}$
2. $\text{m}_{A^*}(x) = \overline{\text{m}_A(x)}$
3. Αν λ είναι ιδιοτιμή του A τότε $\bar{\lambda}$ είναι ιδιοτιμή του A^* .

Άσκηση VII.16 Έστω ένας μη-μηδενικός πίνακας γραμμή $x = (x_1, \dots, x_n) \in \mathbb{F}^{n,1}$ με $n \geq 2$.

Θεωρούμε την ευκλείδεια νόρμα $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$. Θέτουμε

$$A = \mathbb{I}_n + B, \text{ όπου } B = x^t x.$$

Αποδείξτε ότι

1. $B^2 = \|x\|^2 B$ και βρείτε το ελάχιστο πολυώνυμό του.
2. ο A είναι διαγωνοποιήσιμος και βρείτε τις διαστάσεις των ιδιοχώρων του
3. $\det A = 1 + \|x\|^2$.

Άσκηση VII.17 Έστω $A \in \mathbb{C}^{n,n}$ και $a_1, \dots, a_n$ οι γραμμές του A . Ισχύει ότι

$$\|a_1\| \cdots \|a_n\| \geq |\det A|$$

Η ισότητα ισχύει μόνο αν τα $a_1, \dots, a_n$ είναι ορθογώνια ή κάποιο $a_i = 0$.

Άσκηση VII.18 Αποδείξτε ότι αν σε ένα πίνακα $A = (a_{ij})$ ισχύει $|a_{ij}| \leq c$ τότε

$$|\det(A)| \leq c^n n^{n/2}.$$

Η εκτίμηση είναι πολύ καλύτερη από την εκτίμηση $|\det(A)| \leq c^n n!$ που προκύπτει από την πρόταση [III.6.4](#).

Άσκηση VII.19

1. Δίνεται ο πίνακας $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Να βρεθεί ορθογώνιος (δηλαδή πραγματικός μοναδιαίος) $U \in \mathbb{R}^{2 \times 2}$ ώστε ο πίνακας $U^{-1}A$ να είναι άνω τριγωνικός.
2. Να βρεθούν οι ιδιόχωροι και οι ιδιοτιμές του πίνακα $A^{15} - 3A^9 + 15I_2$.

Άσκηση VII.20 Δίνεται ο πίνακας $B = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 2 \\ 0 & 1 & 3 \end{pmatrix}$.

1. Να βρεθεί ορθογώνιος πίνακας $Q \in \mathbb{R}^{3 \times 3}$ ώστε ο πίνακας $Q^{-1}BQ$ να είναι άνω τριγωνικός.
2. Να βρεθεί πολυώνυμο δευτέρου βαθμού $\psi(x) \in \mathbb{R}[x]$ ώστε $BA^{-1} = \psi(B)$.

Άσκηση VII.21 Δίνεται πίνακας $\Gamma \in \mathbb{C}^{n \times n}$. Είναι σωστό ότι ο πίνακας $\Gamma + \Gamma^* + iI_n$ είναι αντιστρέψιμος;

VIII.1 Τετραγωνικές μορφές

VIII.1.1 Τετραγωνικές εξισώσεις στο επίπεδο

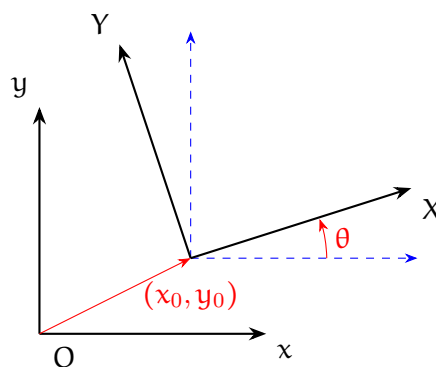
Η γενική δευτεροβάθμια εξίσωση στο επίπεδο έχει την μορφή

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \quad (\text{VIII.1})$$

Σε μαθήματα αναλυτικής γεωμετρίας η παραπάνω εξίσωση μελετάται με την βοήθεια αλλαγής συστήματος συντεταγμένων. Δηλαδή θεωρούμε την αλλαγή μεταβλητών

$$\begin{aligned} x &= x_0 + X \cos \theta - Y \sin \theta \\ y &= y_0 + X \sin \theta + Y \cos \theta \end{aligned} \quad (\text{VIII.2})$$

η οποία οδηγεί σε ένα σύστημα συντεταγμένων μετατοπισμένο κατά (x_0, y_0) και περιστραμμένο κατά γωνία θ όπως στο σχήμα. Στο νέο σύστημα συντεταγμένων η τετραγωνική εξίσωση έχει



Σχήμα VIII.1: Μετακίνηση και περιστροφή συστήματος συντεταγμένων

μετατραπεί στην

$$A'X^2 + 2B'XY + C'Y^2 + 2D'X + 2E'Y + F = 0$$

όπου

$$\begin{aligned} A' &= A \cos^2 \theta + 2B \sin \theta \cos \theta + C \sin^2 \theta \\ B' &= -A \sin \theta \cos \theta + B(\cos^2 \theta - \sin^2 \theta) + C \sin \theta \cos \theta \\ C' &= A \sin^2 \theta - 2B \sin \theta \cos \theta + C \cos^2 \theta \\ D' &= (Ax_0 + By_0 + D) \cos \theta + (Bx_0 + Cy_0 + E) \sin \theta \\ E' &= -(Ax_0 + By_0 + D) \sin \theta + (Bx_0 + Cy_0 + E) \cos \theta \\ F' &= (Ax_0 + By_0 + D)x_0 + B(x_0 + Cy_0 + E)y_0 + Dx_0 + Ey_0 + F \end{aligned}$$

Στην συνέχεια επιλέγουμε το θ ώστε $B' = 0$. Αν $A = C$ τότε αυτό γίνεται αν $\cos 2\theta = 0$ δηλαδή για $\theta = \pi/4$ εκτός αν $B = 0$ οπότε δεν χρειάζεται στροφή. Αν $A \neq B$, τότε θα πρέπει

$$\frac{2B}{A - C} = \frac{2 \sin \theta \cos \theta}{\cos^2 \theta - \sin^2 \theta} = \tan 2\theta$$

από όπου προσδιορίζεται η γωνία θ ώστε $B' = 0$. Στην συνέχεια επιλέγουμε τα (x_0, y_0) ώστε $D', F' = 0$ το οποίο γίνεται λύνοντας το σύστημα

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} D \\ E \end{pmatrix}$$

αν αυτό έχει λύση. Τελικός σκοπός μας είναι να φέρουμε την εξίσωση σε μια μορφή την οποία μπορούμε να αναγνωρίσουμε όπως

$$\begin{aligned} \frac{X^2}{a^2} + \frac{Y^2}{b^2} &= 1 \text{ Έλλειψη} \\ \frac{X^2}{a^2} + \frac{Y^2}{b^2} &= -1 \text{ Κενό σύνολο} \\ \frac{X^2}{a^2} - \frac{Y^2}{b^2} &= 1 \text{ Υπερβολή} \\ -\frac{X^2}{a^2} + \frac{Y^2}{b^2} &= 1 \text{ Υπερβολή} \\ AX^2 + BY^2 &= 0 \text{ Ζευγάρι πραγματικών ευθειών ή σημείο (Μιγαδικές ευθείες)} \end{aligned} \tag{VIII.3}$$

$$A'X^2 + 2D'X + 2E'Y + F' = 0 \text{ Παραβολή ή ζεύγος ευθειών πραγματικών ή μιγαδικών}$$

$$C'Y^2 + 2D'X + 2E'Y + F' = 0 \text{ Παραβολή ή ζεύγος ευθειών πραγματικών ή μιγαδικών}$$

(VIII.4)

Η μέθοδος αυτή έχει το πλεονέκτημα της γεωμετρικής εποπτίας αλλά απαιτεί αρκετές πράξεις και γενικεύεται δύσκολα σε μεγαλύτερες διαστάσεις. Επιπλέον η διερεύνηση είναι αρκετά πολύπλοκη.

Πριν προχωρήσουμε σε μία μέθοδο βασισμένη στις τεχνικές του μαθήματος αυτού ας παρατηρήσουμε ότι η αλλαγή συστήματος συντεταγμένων στην εξίσωση (VIII.2) μπορεί να γραφεί σε μορφή πινάκων ως

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \text{ με } P = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \tag{VIII.5}$$

όπου ο πίνακας P είναι μοναδιαίος. Επίσης παρατηρούμε ότι η εξίσωση (VIII.1) γράφεται στην μορφή

$$(x, y) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + 2 \begin{pmatrix} E & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + F = 0 \tag{VIII.6}$$

Δηλαδή, ο υπολογισμός της γωνίας θ αντιστοιχεί στον υπολογισμό ενός μοναδιαίου πίνακα P ώστε

$$P^t \begin{pmatrix} A & B \\ B & C \end{pmatrix} P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

με $\lambda_1, \lambda_2 \in \mathbb{R}$ ώστε αντικαθιστώντας την (VIII.5) στην (VIII.6) έχουμε ότι

$$\begin{aligned} (x, y) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= (X \ Y) P^t \begin{pmatrix} A & B \\ B & C \end{pmatrix} P \begin{pmatrix} X \\ Y \end{pmatrix} + (x_0 \ y_0) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\ &= (X \ Y) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + (x_0 \ y_0) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\ &= \lambda_1 X^2 + \lambda_2 Y^2 + (x_0 \ y_0) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \end{aligned}$$

Στην παραπάνω εξίσωση ο πίνακας P είναι αντιστρέψιμος οπότε θέτουμε

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = P \begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix}$$

και έχουμε ότι

$$\begin{aligned} (x, y) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \lambda_1 X^2 + \lambda_2 Y^2 + (x_0 \ y_0) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\ &= \lambda_1 X^2 + \lambda_2 Y^2 + \lambda_1 (x'_0)^2 + \lambda_2 (y'_0)^2 \\ &= \lambda_1 (X - x'_0)^2 + \lambda_2 (Y - y'_0)^2 + 2(x'_0 \ y'_0) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \end{aligned}$$

Από την άλλη έχουμε

$$2(E \ D) \begin{pmatrix} x \\ y \end{pmatrix} = 2(E \ D) P \begin{pmatrix} X \\ Y \end{pmatrix} + 2(E \ D) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

Οπότε αν μπορέσουμε να διαλέξουμε το (x_0, y_0) με τέτοιο τρόπο ώστε

$$(x'_0 \ y'_0) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + (E \ D) P = 0$$

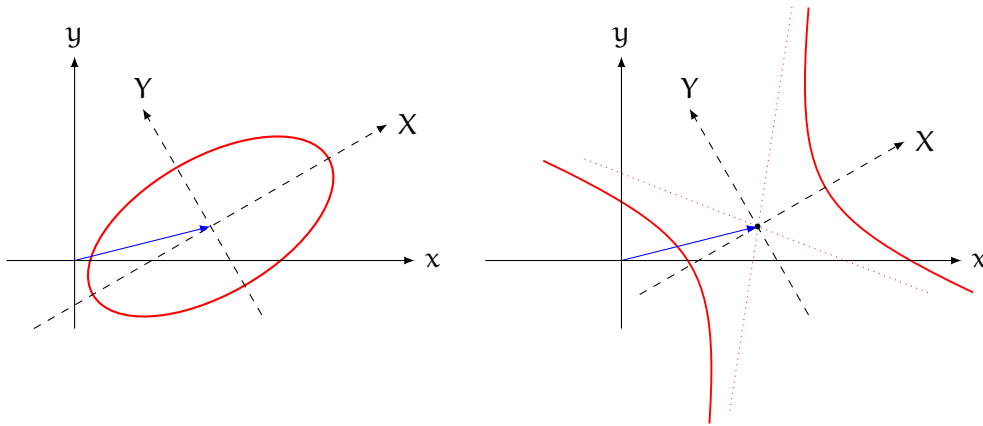
ισοδύναμα αν το σύστημα

$$A^t \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = - \begin{pmatrix} E \\ D \end{pmatrix} \quad (\text{VIII.7})$$

έχει λύση $(x_0, y_0)^t$, τότε η αρχική εξίσωση γράφεται στην μορφή

$$\lambda_1 (X - x'_0)^2 + \lambda_2 (Y - y'_0)^2 + 2(E \ D) P \begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix} + F = 0.$$

Αυτή την περίπτωση την λέμε καμπύλη με κέντρο συμμετρίας το (x_0, y_0) . Το παραπάνω σύστημα έχει μοναδική λύση αν και μόνο αν η ποσότητα $AC - B^2 = \lambda_1 \lambda_2 \neq 0$. Οι ιδιοτιμές είναι πραγματικές, ως ιδιοτιμές συμμετρικού πίνακα. Καταλλήγουμε με αυτό τον τρόπο στις πέντε πρώτες εξισώσεις της ταξινόμησης στην εξίσωση (VIII.3). Στο παρακάτω σχήμα σχεδιάζουμε μία περιστραμμένη έλλειψη και μία περιστραμμένη υπερβολή.



Αν υποθέσουμε τώρα ότι το σύστημα (VIII.7) δεν έχει λύση. Αυτό σημαίνει ότι μία τουλάχιστον ιδιοτιμή είναι μηδενική. Υποθέτουμε ότι $\lambda_1 \neq 0$ και $\lambda_2 = 0$. Η εξίσωση της καμπύλης γίνεται σε αυτή την περίπτωση

$$\lambda_1(X - x'_0)^2 + 2 \begin{pmatrix} x'_0 & y'_0 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + 2(E \ D)P \begin{pmatrix} X \\ Y \end{pmatrix} + 2(E \ D) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + F = 0.$$

Σε αυτή την περίπτωση διαλέγουμε x'_0 ώστε

$$x'_0 \lambda_1 + Ep_{11} + Dp_{22} = 0$$

οπότε η εξίσωση γίνεται

$$\lambda_1(X - x'_0)^2 + 2(Ep_{12} + Dp_{22})Y + 2(E \ D)P \begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix} + F = 0$$

Αν $Ep_{12} + Dp_{22} \neq 0$ η παραπάνω εξίσωση παριστάνει μία παραβολή ενώ αν $Ep_{12} + Dp_{22} = 0$ η εξίσωση είναι δύο ευθείες πραγματικές ή μιγαδικές.

Αν και οι δύο ιδιοτιμές $\lambda_1 = \lambda_2 = 0$ είναι μηδενικές τότε η καμπύλη δεν είναι τετραγωνική αλλά εκφυλίζεται στην τόπο μηδενισμού της

$$2Dx + 2Ey + F = 0$$

η οποία είναι ευθεία ή αδύνατη (αν $D = E = 0$ και $F \neq 0$) ή όλο το $\mathbb{R}^2$ αν $D = E = F = 0$.

VIII.1.2 Τετραγωνικές εξισώσεις στον χώρο

Η γενική μορφή μιας τετραγωνικής εξίσωσης στον χώρο είναι η

$$(x_1, \dots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + 2(E_1 \ E_2 \ \cdots \ E_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + F = 0.$$

Θέτουμε $x = (x_1, \dots, x_n)^t$, $E = (E_1, \dots, E_n)$ και $A = (a_{ij})$ είναι ένας συμμετρικός πίνακας. Συνεπώς υπάρχει μοναδιαίος πίνακας $P \in \mathbb{R}^{n,n}$ ώστε

$$P^t A P = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Θέτουμε $X = (X_1, \dots, X_n)$ από την εξίσωση $X = P^t x$ δηλαδή $x = PX$. Έχουμε την εξίσωση

$$X^t P^t A P X + 2EPX + F = X^t \text{diag}(\lambda_1, \dots, \lambda_n) X + 2EPX + F = 0.$$

Η παραπάνω εξίσωση δίνει

$$\sum_{i=1}^n \lambda_i X_i^2 + 2 \sum_{i=1}^n \sum_{v=1}^n E_v P_{v,i} X_i + F = 0.$$

Υποθέτουμε ότι οι ιδιοτιμές $\lambda_1, \dots, \lambda_\kappa$ είναι διαφορετικές του μηδενός ενώ $\lambda_{\kappa+1} = \dots = \lambda_n = 0$. Για κάθε ιδιοτιμή λ_i , $1 \leq i \leq \kappa$ γράφουμε $Y_i = X_i + a_i$ και έχουμε ότι

$$\begin{aligned} \lambda_i (X_i + a_i - a_i)^2 + 2 \sum_{v=1}^n E_v P_{v,i} (X_i + a_i - a_i) &= \lambda_i Y_i^2 - 2a_i \lambda_i Y_i + \lambda_i a_i^2 + 2 \sum_{v=1}^n E_v P_{v,i} Y_i - 2 \sum_{v=1}^n E_v P_{v,i} a_i \\ &= \lambda_i Y_i^2 - \frac{1}{\lambda_i} \left(\sum_{v=1}^n E_v P_{v,i} \right)^2 \end{aligned} \quad (\text{VIII.8})$$

επιλέγοντας

$$a_i = \frac{1}{\lambda_i} \sum_{v=1}^n E_v P_{v,i},$$

ώστε να μηδενιστεί ο συντελεστής του Y_i .

Με αυτό τον τρόπο καταλήγουμε στην εξίσωση

$$\sum_{i=1}^{\kappa} \lambda_i Y_i^2 + \sum_{i=\kappa+1}^n Q_i X_i - \sum_{i=1}^{\kappa} \frac{1}{\lambda_i} Q_i^2 + F = 0 \quad (\text{VIII.9})$$

όπου

$$Q_i = \sum_{v=1}^n E_v P_{v,i}.$$

Αν $\kappa = n$ ή αν $Q_{\kappa+1} = \dots = Q_n = 0$ τότε η εξίσωση γράφεται ως

$$\sum_{i=1}^{\kappa} \lambda_i Y_i^2 + F = 0$$

η οποία ονομάζεται κανονική εξίσωση τετραγωνικής υπερεπιφάνειας με κέντρο. Αν $\kappa = n$ τότε ονομάζεται μη εκφυλισμένη διαφορετικά ονομάζεται εκφυλισμένη.

Η εξίσωση (VIII.9) μπορεί να απλοποιηθεί περαιτέρω αν θέσουμε ως νέα μεταβλητή την

$$Y = \sum_{i=\kappa+1}^n Q_i X_i.$$

Προκειμένου να έχουμε διατήρηση του εσωτερικού γινομένου θέλουμε οι αλλαγές μεταβλητών να γίνονται μέσω μοναδιαίων πινάκων. Οπότε θεωρούμε τον πίνακα

$$P_1 = \left(\begin{array}{cccc|ccc} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & \dots & 0 & -\frac{1}{q} Q_{\kappa+1} & \dots & -\frac{1}{q} Q_n \\ q_{\kappa+2,1} & q_{\kappa+2,2} & \dots & \dots & \dots & \dots & q_{\kappa+1,n} \\ \vdots & \vdots & & & & & \vdots \\ q_{\kappa+1,1} & q_{\kappa+1,2} & \dots & \dots & \dots & \dots & q_{\kappa+1,n} \end{array} \right)$$

στον οποίο το πάνω δεξιά $\kappa \times \kappa$ κομμάτι είναι ο ταυτοτικός πίνακας, η $\kappa+1$ γραμμή έχει διαιρεθεί με το μέτρο της

$$q = \sqrt{\sum_{v=\kappa+1}^n Q_v^2}$$

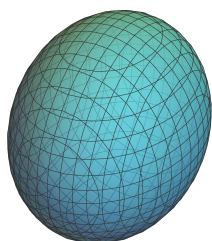
ώστε οι πρώτες $\kappa + 1$ γραμμές να είναι ορθοκανονικές ενώ έχουμε συμπληρώσει τις υπόλοιπες γραμμές σε μία ορθοκανονική βάση του $\mathbb{R}^n$ και ο πίνακας P_1 να είναι μοναδιαίος. Εκτελούμε μια αλλαγή βάσης $Y = P_1^t Z$ οπότε καταλήγουμε στην εξίσωση

$$\sum_{i=1}^{\kappa} \lambda_i Z_i^2 = 2qZ_{\kappa+1} - D.$$

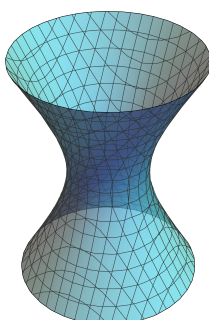
Αν $D \neq 0$ τότε με μια παράλληλη μετατόπιση $T = Z_{\kappa+1} - \frac{F}{2q}$ έχουμε την εξίσωση

$$\sum_{i=1}^{\kappa} \lambda_i Z_i^2 = 2qT,$$

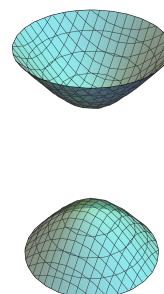
η οποία ονομάζεται η κανονική εξίσωση μιας υπερεπιφάνειας χωρίς κέντρο.



$$(α') \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

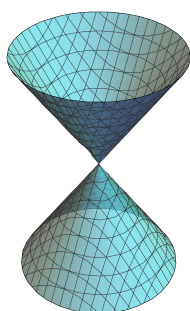


$$(β') \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

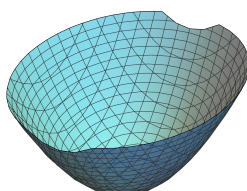


$$(γ') \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

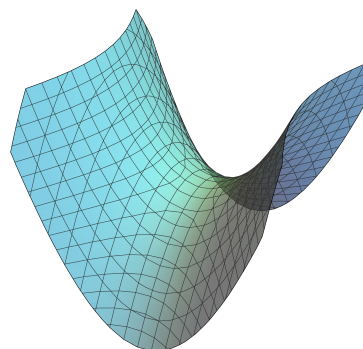
Σχήμα VIII.2: $\kappa = 3$ μη-ιδιόμορφες



$$(α'') \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$$



$$(β'') \frac{x^2}{a^2} + \frac{y^2}{b^2} = z$$



$$(γ'') \frac{x^2}{a^2} - \frac{y^2}{b^2} = z$$

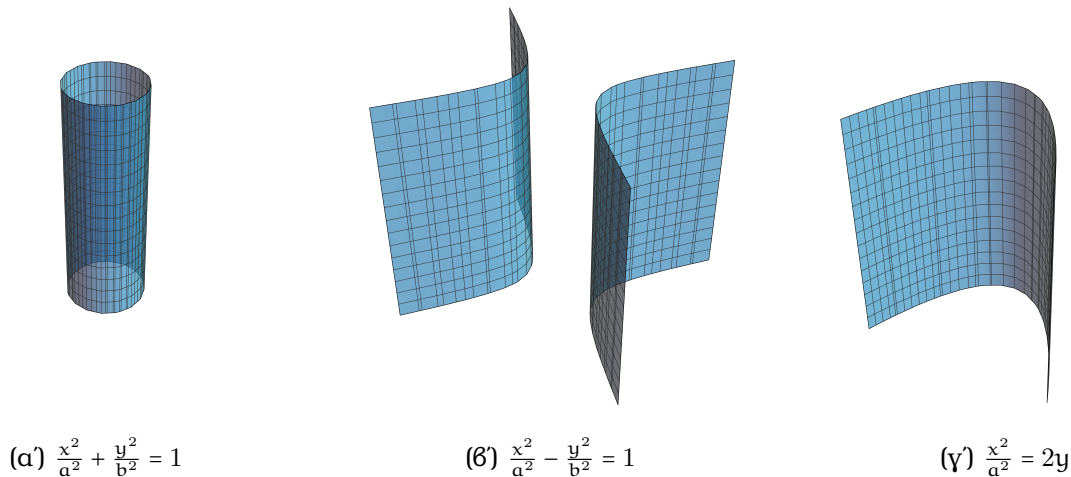
Σχήμα VIII.3: $\kappa = 3$ μη-ιδιόμορφες

Παράδειγμα VIII.1.1. Για να μελετήσουμε την επιφάνεια

$$x^2 - 2xy + 2y^2 = 1,$$

την γράφουμε στην μορφή

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1.$$

Σχήμα VIII.4: $\kappa = 1$ κύλινδροι

Υπολογίζουμε τις ιδιοτιμές $\lambda_1 = \frac{3-\sqrt{5}}{2}$, $\lambda_2 = \frac{3+\sqrt{5}}{2}$ και βρίσκουμε τους αντίστοιχους ιδιόχωρους $E_{\lambda_1} = \langle (\frac{1+\sqrt{5}}{2}, 1)^t \rangle$, $E_{\lambda_2} = \langle (\frac{1-\sqrt{5}}{2}, 1)^t \rangle$. Τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους γιατί αντιστοιχούν σε ιδιοτιμές συμμετρικού πίνακα. Για να καταλήξουμε σε μία ορθοκανονική βάση τα διαιρούμε με το μέτρο τους και έχουμε τα ορθοκανονικά διανύσματα

$$e_1 = \frac{(\frac{1+\sqrt{5}}{2}, 1)^t}{\|(\frac{1+\sqrt{5}}{2}, 1)^t\|} = \left(\frac{1+\sqrt{5}}{\sqrt{2(5+\sqrt{5})}}, \sqrt{\frac{2}{5+\sqrt{5}}} \right)^t \quad e_2 = \frac{(\frac{1-\sqrt{5}}{2}, 1)^t}{\|(\frac{1-\sqrt{5}}{2}, 1)^t\|} = \left(-\frac{\sqrt{5}-1}{\sqrt{2(5-\sqrt{5})}}, \sqrt{\frac{2}{5-\sqrt{5}}} \right)^t$$

τα οποία τα τοποθετούμε ως στήλες του μοναδιαίου πίνακα P

$$P = \begin{pmatrix} \frac{1+\sqrt{5}}{\sqrt{2(5+\sqrt{5})}} & -\frac{\sqrt{5}-1}{\sqrt{2(5-\sqrt{5})}} \\ \sqrt{\frac{2}{5+\sqrt{5}}} & \sqrt{\frac{2}{5-\sqrt{5}}} \end{pmatrix}.$$

Θέτοντας ως , από τον τύπο

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix}$$

έχουμε την εξίσωση στην μορφή

$$\lambda_1 X^2 + \lambda_2 Y^2 = 1$$

η οποία παριστάνει μια έλλειψη αφού και οι δύο ιδιοτιμές λ_1, λ_2 είναι θετικές, όπως φαίνεται στο σχήμα **VIII.5**. Στην συνέχεια θέλουμε να μελετήσουμε την μορφή της καμπύλης

$$x^2 - 2xy + 2y^2 + x + y + F = 0$$

για διάφορες τιμές της παραμέτρου $F \in \mathbb{R}$. Υπολογίζουμε ότι

$$(Q_1 \ Q_2) = (E_1 \ E_2) P = \left(\frac{3+\sqrt{5}}{\sqrt{2(5+\sqrt{5})}} \quad \frac{3-\sqrt{5}}{\sqrt{2(5-\sqrt{5})}} \right).$$

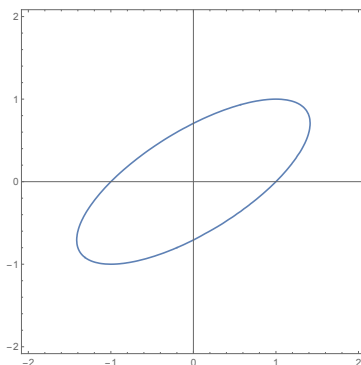
Υπολογίζουμε ότι

$$a_1 = Q_1/\lambda_1 = -\frac{(3+\sqrt{5})\sqrt{\frac{2}{5+\sqrt{5}}}}{\sqrt{5}-3} \quad a_2 = Q_2/\lambda_2 = -\frac{(\sqrt{5}-3)\sqrt{\frac{1}{10}(5+\sqrt{5})}}{3+\sqrt{5}},$$

οπότε για τις μεταβλητές $Y_i = X_i + a_i$, $i = 1, 2$, σύμφωνα με την εξίσωση **(VIII.8)** έχουμε την εξίσωση

$$\lambda_1 Y_1^2 + \lambda_2 Y_2^2 = -F + \frac{Q_1^2}{\lambda_1} + \frac{Q_2^2}{\lambda_2} = 5 - F.$$

Οπότε η εξίσωση παριστάνει όπως και πριν έλλειψη αρκεί $5 - F > 0$ ή σημείο αν $F = 5$ ή το κενό σύνολο αν $5 - F < 0$.

Σχήμα VIII.5: Η έλλειψη $x^2 - 2xy + 2y^2 = 0$

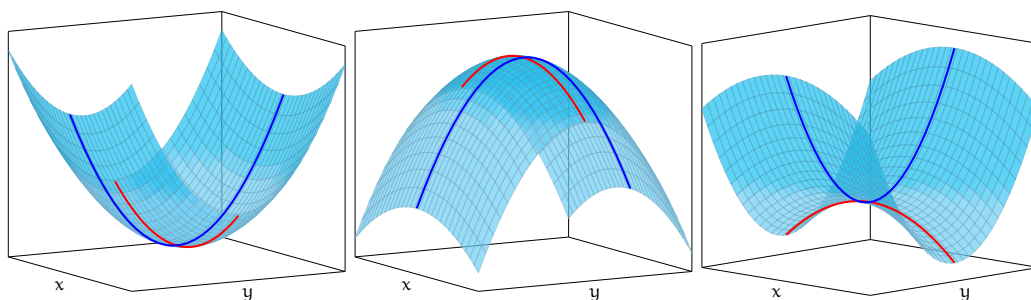
VIII.1.3 Μελέτη ακροτάτων

Όπως είδαμε στην πρόταση VII.2.10 ένα σημείο x_0 είναι τοπικό ακρότατο της f αν το x_0 είναι κρίσιμο σημείο δηλαδή $Df(x_0) = 0$. Στην περίπτωση αυτή η έκφραση (VII.6) γράφεται ως

$$f(x) = f(x_0) + (x - x_0)^t H(x_0)(x - x_0) + \text{όροι μεγαλύτερης τάξης},$$

και η συμπεριφορά της f καθορίζεται από την τετραγωνική εξίσωση που παραλείπει τους όρους μεγαλύτερης τάξης αν αυτοί υπάρχουν. Πράγματι όταν η ποσότητα $x - x_0$ είναι μικρή, οι κυβικοί και οι μεγαλύτεροι όροι επηρεάζουν λιγότερο την $f(x)$.

Η συμπεριφορά της τετραγωνικής εξίσωσης καθορίζεται από της ιδιοτιμές του Hessian πίνακα ο οποίος είναι συμμετρικός και συνεπώς είναι διαγωνοποιήσιμος με πραγματικές ιδιοτιμές. Παρακάτω έχουμε σχεδιάσει τις περιπτώσεις για τις ιδιοτιμές όταν ο πίνακας $H(x_0)$ είναι αντιστρέψιμος 2×2 .



Στην γενική περίπτωση για το $n \geq 1$ ο πίνακας $H(x_0)$ έχει πραγματικές ιδιοτιμές ως συμμετρικός. Είναι σαφές ότι ο περιορισμός της f πάνω στον ιδιόχωρο της ιδιοτιμής λ_i είναι η παραβολή $f(\xi) = \lambda_i \xi^2$ η οποία έχει ελάχιστο στο κρίσιμο σημείο, όταν η ιδιοτιμή είναι θετική και μέγιστο όταν η ιδιοτιμή είναι αρνητική. Αν δηλαδή ο πίνακας $H(x_0)$ έχει όλες τις ιδιοτιμές θετικές, είναι δηλαδή θετικά ορισμένος τότε το κρίσιμο σημείο είναι ελάχιστο ενώ αν έχει όλες τις ιδιοτιμές αρνητικές, δηλαδή ο $-H(x_0)$ είναι θετικά ορισμένος τότε το κρίσιμο σημείο είναι μέγιστο. Στην περίπτωση που ο πίνακας έχει και θετικές και αρνητικές ιδιοτιμές τότε ο περιορισμός της f στον χώρο που παράγεται από τα ιδιοδιανύσματα που αντιστοιχούν σε θετικές ιδιοτιμές έχει ελάχιστο ενώ ο περιορισμός στον χώρο που παράγεται από τα ιδιοδιανύσματα που αντιστοιχούν σε αρνητικές ιδιοτιμές έχει μέγιστο. Γενικά το κρίσιμο σημείο δεν είναι ούτε μέγιστο ούτε ελάχιστο αλλά είναι ένα σαγματικό στοιχείο και αντιστοιχεί στο δεξί γράφημα στο παραπάνω σχήμα στην περίπτωση $n = 2$. Τέλος αν περιορίσουμε την f στον πυρήνα της $H(x_0)$ δηλαδή στον ιδιόχωρο της ιδιοτιμής 0, δεν μπορούμε να έχουμε κανένα συμπέρασμα για την f . Η συμπεριφορά της εξαρτάται από τους όρους μεγαλύτερης τάξης.

VIII.2 Μιάμιση και διγραμμικές μορφές

Ορισμός VIII.2.1. Έστω V ένας διανυσματικός χώρος πάνω από το $\mathbb{F}$, όπου $\mathbb{F} = \mathbb{R}$ ή $\mathbb{F} = \mathbb{C}$. Μία συνάρτηση

$$f : V \times V \longrightarrow \mathbb{F}$$

θα λέγεται *μιάμιση (sesqui) μορφή* αν ικανοποιεί

1. $f(\lambda_1 v_1 + \lambda_2 v_2, v_3) = \lambda_1 f(v_1, v_3) + \lambda_2 f(v_2, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.
2. $f(v_1, \lambda_1 v_2 + \lambda_2 v_3) = \bar{\lambda}_1 f(v_1, v_2) + \bar{\lambda}_2 f(v_1, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.

Θα λέγεται *διγραμμική* αν ικανοποιεί τις σχέσεις

1. $f(\lambda_1 v_1 + \lambda_2 v_2, v_3) = \lambda_1 f(v_1, v_3) + \lambda_2 f(v_2, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.
2. $f(v_1, \lambda_1 v_2 + \lambda_2 v_3) = \lambda_1 f(v_1, v_2) + \lambda_2 f(v_1, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.

Παρατήρηση VIII.2.2. Αν το $\mathbb{F} = \mathbb{R}$ τότε οι έννοιες της διγραμμικής και μιάμιση μορφής ταυτίζονται.

Παράδειγμα VIII.2.3. Σε ένα διανυσματικό χώρο V υπέρ του σώματος $\mathbb{C}$ ένα ερμητιανό εσωτερικό γινόμενο είναι μια μιάμιση μορφή. Επίσης αν $L : V \rightarrow V$ είναι μια γραμμική συνάρτηση τότε η

$$f(v, w) = \langle Lv, w \rangle$$

είναι μια μιάμιση μορφή.

Σχόλιο: Μία διγραμμική συνάρτηση είναι διπλά γραμμική. Αντίθετα μια μιάμιση μορφή είναι γραμμική στο πρώτο όρισμα αλλά κάτι λιγότερο από γραμμική στο δεύτερο αφού ισχύει μόνο η προσθετική ιδιότητα της γραμμικής συνάρτησης. Η λατινική λέξη “sesqui” σημαίνει ένα και μισό.

VIII.2.1 Μιάμιση μορφές

Θεώρημα VIII.2.4. Για ένα διανυσματικό χώρο V πεπερασμένης διάστασης με εσωτερικό γινόμενο και μια μιάμιση μορφή f υπάρχει μοναδική γραμμική συνάρτηση $L_f : V \rightarrow V$ ώστε

$$f(v, w) = \langle L_f v, w \rangle \text{ για κάθε } v, w \in V.$$

Απόδειξη. Για ένα σταθερό διάνυσμα w έχουμε ότι η συνάρτηση

$$\begin{aligned} f(v, w) : V &\longrightarrow \mathbb{F} \\ v &\longmapsto f(v, w) \end{aligned}$$

είναι μια γραμμική συνάρτηση και συνεπώς από την πρόταση [VII.7.3](#) έχουμε ότι υπάρχει ένα $u(w) \in V$ ώστε

$$f(v, w) = \langle v, u(w) \rangle.$$

Στην παραπάνω εξίσωση συμβολίσαμε το στοιχείο με $u(w)$ για να δείξουμε την εξάρτησή του από το w . Θα δείξουμε τώρα ότι

$$u(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 u(w_1) + \lambda_2 u(w_2) \quad (\text{VIII.10})$$

δηλαδή το στοιχείο $u(w)$ είναι γραμμική συνάρτηση του w . Πράγματι έχουμε για κάθε $v, w_1, w_2 \in V$ και κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ ότι

$$\begin{aligned}\langle v, u(\lambda_1 w_1 + \lambda_2 w_2) \rangle &= f(v, \lambda_1 w_1 + \lambda_2 w_2) \\ &= \bar{\lambda}_1 f(v, w_1) + \bar{\lambda}_2 f(v, w_2) \\ &= \langle v, \lambda_1 u(w_1) \rangle + \langle v, \lambda_2 u(w_2) \rangle \\ &= \langle v, \lambda_1 u(w_1) + \lambda_2 u(w_2) \rangle\end{aligned}$$

από όπου προκύπτει ότι ισχύει η **(VIII.10)**. Έτσι αρκεί να πάρουμε $L_f = u^*$. Τέλος παρατηρούμε ότι αν υπήρχαν δύο συναρτήσεις L_1, L_2 οι οποίες να είχαν την ζητούμενη ιδιότητα τότε θα είχαμε ότι για κάθε $v, w \in V$

$$\langle L_1(v), w \rangle = \langle L_2(v), w \rangle \Rightarrow \langle L_1(v) - L_2(v), w \rangle = 0$$

συνεπώς $L_1 = L_2$. □

Έστω F ο χώρος των μιάνιση μορφών είναι ένας διανυσματικός χώρος πάνω από το $\mathbb{F}$. Πράγματι αν προσθέσουμε δύο μιάνιση μορφές παίρνουμε μια μιάνιση μορφή και αν πολλαπλασιάσουμε μια μιάνιση μορφή με ένα στοιχείο $\lambda \in \mathbb{F}$ παίρνουμε μια μιάνιση μορφή.

Πρόταση VIII.2.5. Υπάρχει ένας ισομορφισμός διανυσματικών χώρων

$$\begin{aligned}\psi : F &\longrightarrow \text{End}(V) \\ f &\longmapsto L_f\end{aligned}$$

Απόδειξη. Παρατηρούμε ότι για οποιαδήποτε στοιχεία $f_1, f_2 \in F$ και οποιαδήποτε $\lambda_1, \lambda_2 \in \mathbb{F}$ έχουμε ότι

$$\begin{aligned}\langle L_{\lambda_1 f_1 + \lambda_2 f_2}(v), w \rangle &= (\lambda_1 f_1 + \lambda_2 f_2)(v, w) = \lambda_1 f_1(v, w) + \lambda_2 f_2(v, w) \\ &= \lambda_1 \langle L_{f_1}(v), w \rangle + \lambda_2 \langle L_{f_2}(v), w \rangle = \langle (\lambda_1 L_{f_1} + \lambda_2 L_{f_2})(v), w \rangle\end{aligned}$$

από όπου προκύπτει ότι

$$L_{\lambda_1 f_1 + \lambda_2 f_2} = \lambda_1 L_{f_1} + \lambda_2 L_{f_2}$$

δηλαδή η συνάρτηση ψ είναι γραμμική. Επίσης κάθε $L : V \rightarrow V$ ορίζει μια $f(v, w) = \langle Lv, w \rangle$ άρα η ψ είναι επί. Τέλος $\ker \psi$ είναι οι $f \in F$ ώστε $L_f = 0$ αλλά τότε $f(v, w) = \langle L_f v, w \rangle = \langle 0, w \rangle = 0$, δηλαδή η ψ είναι και 1-1. □

Ορισμός VIII.2.6. Για μια μιάνιση μορφή και για μία διατεταγμένη βάση $B = \{v_1, \dots, v_n\}$ του V , ορίζουμε τον πίνακα $A = (a_{ij})$ με

$$a_{ij} = f(v_j, f_i),$$

τον οποίο ονομάζουμε πίνακα της μιάνιση μορφής f .

Ο πίνακας A της μιάνιση μορφής γνωρίζει ότι την πληροφορία της f , όπως δείχνει η παρακάτω πρόταση.

Πρόταση VIII.2.7. Για v, w συμβολίζουμε με $[v]_B, [w]_B \in \mathbb{R}^n$ τα διανύσματα στήλτες των συ-

υτεταγμένων των v, w . Ισχύει

$$f(v, w) = [w]_B^* A [v]_B.$$

Απόδειξη. Γράφουμε $[v]_B = (\lambda_1, \dots, \lambda_n)^t$ και $[w]_B = (\mu_1, \dots, \mu_n)^t$. Έχουμε ότι

$$\begin{aligned} f(v, w) &= f\left(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \bar{\mu}_j \lambda_i f(v_i, v_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{j,i} \bar{\mu}_j \lambda_i \end{aligned}$$

από όπου προκύπτει το ζητούμενο. $\square$

Παρατήρηση VIII.2.8. Αν επιλέξουμε την βάση $v_1, \dots, v_n$ να είναι ορθοκανονική τότε έχουμε ότι

$$f(v_j, v_i) = \langle T_f e_j, e_i \rangle$$

ο οποίος είναι ο πίνακας της γραμμικής συνάρτησης T_f ως προς την βάση B , δείτε την πρόταση VII.7.9.

Παρατήρηση VIII.2.9. Αν κάνουμε την αλλαγή βάσης $v'_j = \sum_{i=1}^n p_{ij} v_i$ από την βάση B στην βάση $B' = \{v'_1, \dots, v'_n\}$ τότε έχουμε ότι ο πίνακας A' της μιάμιση μορφής ως προς την νέα βάση δίνεται από την εξίσωση

$$A' = P^* A P.$$

Πράγματι,

$$\begin{aligned} a'_{ij} &= f(v'_j, v'_i) = f\left(\sum_{\nu=1}^n p_{\nu,j} v_\nu, \sum_{\mu=1}^n p_{\mu,i} v_\mu\right) \\ &= \sum_{\nu=1}^n \sum_{\mu=1}^n p_{\nu,j} \bar{p}_{\mu,i} f(v_\nu, v_\mu) \\ &= \sum_{\nu=1}^n \sum_{\mu=1}^n \bar{p}_{\mu,i} a_{\mu,\nu} p_{\nu,j} \end{aligned}$$

Ορισμός VIII.2.10. Μια μιάμιση μορφή f θα λέγεται ερμητιανή αν και μόνο αν

$$f(v, w) = \overline{f(w, v)} \text{ για κάθε } v, w \in V.$$

Παρατήρηση VIII.2.11. Αν $L_f : V \rightarrow V$ είναι η γραμμική συνάρτηση που αντιστοιχεί στην μιάμιση μορφή f , έχουμε ότι

$$\begin{aligned} f(v, w) &= \langle L_f(v), w \rangle \\ \overline{f(w, v)} &= \langle L_f(w), v \rangle = \langle w, L_f^*(v) \rangle \end{aligned}$$

από όπου προκύπτει ότι η f είναι ερμητιανή αν και μόνο αν η L_f είναι ερμητιανή, δηλαδή $L_f^* = L_f$.

Θεώρημα VIII.2.12. Για κάθε ερμητιανή μιάμιση μορφή f σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V με εσωτερικό γινόμενο υπάρχει μια ορθοκανονική βάση στην οποία ο f αναπαρίσταιται από ένα διαγώνιο πίνακα με πραγματικά στοιχεία στην διαγώνιο.

Απόδειξη. Αφού η f είναι ερμητιανή, σύμφωνα με την παρατήρηση VIII.2.11 η γραμμική της συνάρτηση L_f είναι ερμητιανή, δηλαδή $L_f = L_f^*$. Σύμφωνα με το θεώρημα VII.9.6 υπάρχει ορθοκανονική βάση $B = \{e_1, \dots, e_n\}$ του V η οποία αποτελείται από ιδιοδιανύσματα της L_f , δηλαδή $L_f(e_i) = \lambda_i e_i$. Έχουμε ότι

$$f(e_i, e_j) = \langle L_f(e_i), e_j \rangle = \lambda_i \delta_{ij}.$$

Τα $\lambda_i \in \mathbb{R}$ σύμφωνα με την πρόταση VII.9.4. □

Ορισμός VIII.2.13. Μία μιάμιση μορφή f θα λέγεται μη-αρνητική αν και μόνο αν είναι ερμητιανή και επιπλέον $f(v, v) \geq 0$ για κάθε $v \in V$. Η μιάμιση μορφή θα λέγεται θετική αν είναι ερμητιανή και επιπλέον $f(v, v) > 0$ για κάθε $v \in V, v \neq 0$.

Παρατήρηση VIII.2.14. Απαιτούμε η f να είναι ερμητιανή γιατί τότε $f(v, v) \in \mathbb{R}$, δείτε την άσκηση VIII.3, οπότε η ανισότητα έχει νόημα.

Παρατήρηση VIII.2.15. Μια θετική μιάμιση μορφή είναι ένα εσωτερικό γινόμενο στον V . Μια μη-αρνητική έχει όλες τις ιδιότητες εκτός από το ότι μπορεί να υπάρχουν και μη-μηδενικά $v \in V$ με $f(v, v) = 0$.

Παρατήρηση VIII.2.16. Αν $A = (a_{ij})$ με $a_{ij} = f(v_j, v_i)$ είναι ο πίνακας της f ως προς μία διατεταγμένη βάση $B = \{v_1, \dots, v_n\}$ του V και $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in V$ τότε

$$f(v, v) = [v]_B^* A [v]_B = \sum_{j=1}^n \sum_{i=1}^n a_{ij} \lambda_j \bar{\lambda}_i. \quad (\text{VIII.11})$$

Συνεπώς για να είναι ο A μη-αρνητικός θα πρέπει $A = A^*$ και επιπλέον θα πρέπει η έκφραση της εξίσωσης (VIII.11) να είναι θετική για κάθε $(\lambda_1, \dots, \lambda_n)^t \in \mathbb{R}^n$.

Πρόταση VIII.2.17. Έστω $A \in \mathbb{F}^{n,n}$. Η συνάρτηση

$$\begin{aligned} \mathbb{F}^n \times \mathbb{F}^n &\longrightarrow \mathbb{F} \\ (x, y) &\longmapsto f(x, y) = y^* A x, \end{aligned}$$

δίνει μια θετική μορφή αν και μόνο αν υπάρχει αντιστρέψιμος $P \in \mathbb{F}^{n,n}$ ώστε $A = P^* P$.

Απόδειξη. Παρατηρούμε ότι αν $A = P^* P$ τότε

$$f(x, x) = x^* A x = x^* P^* P x = (Px)^* P x \geq 0.$$

Επιπλέον αν $(Px)^* P x = 0$ τότε $Px = 0$ και αν ο P είναι αντιστρέψιμος τότε $x = 0$.

Αντιστρόφως, αν η $f(x, y)$ είναι θετική τότε είναι εσωτερικό γινόμενο και μπορούμε να διαλέξουμε μία ορθοκανονική βάση $Q_1, \dots, Q_n \in \mathbb{F}^n$, ώστε

$$f(Q_i, Q_j) = Q_j^* A Q_i = \delta_{ij}.$$

Το τελευταίο έχει ως συνέπεια ότι ο πίνακας P που έχει ως στήλες τα $Q_1, \dots, Q_n$ ικανοποιεί

$$Q^* A Q = I_n \Rightarrow A = (Q^{-1})^* Q,$$

δηλαδή το ζητούμενο, αφού ο P είναι αντιστρέψιμος. □

Ορισμός VIII.2.18. Για ένα πίνακα $A = (a_{ij})$ οι κύριες ελάσσονες ορίζουσες $\Delta_k(A)$ είναι οι ορίζουσες

$$\Delta_k(A) = \det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \quad 1 \leq k \leq n.$$

Λήμμα VIII.2.19. Τα παρακάτω είναι ισοδύναμα για $A \in \mathbb{F}^{n,n}$

1. Υπάρχει άνω τριγωνικός πίνακας P με μονάδες στην διαγώνιο ώστε ο πίνακας $B = AP$ να είναι κάτω τριγωνικός.
2. Τα $\Delta_k(A) \neq 0$ για κάθε $1 \leq k \leq n$.

Απόδειξη. Για ένα πίνακα $P = (p_{ij})$ και $A = (a_{ij})$ έχουμε

$$b_{ij} = \sum_{v=1}^n a_{iv} p_{vj}.$$

Αν ο πίνακας P είναι άνω τριγωνικός, δηλαδή $p_{v,j} = 0$ για $v > j$ και $p_{jj} = 1$ τότε η παραπάνω σχέση δίνει:

$$\sum_{v=1}^{j-1} a_{iv} p_{v,j} = b_{ij} - a_{ij} \quad \text{για κάθε } 1 \leq j \leq n.$$

Για να είναι ο πίνακας $B = (b_{ij})$ κάτω τριγωνικός θα πρέπει $b_{ij} = 0$ για $i < j$ το οποίο συμβαίνει αν και μόνο αν

$$\sum_{v=1}^{j-1} a_{iv} p_{v,j} = -a_{ij}$$

για όλα τα $1 \leq i \leq j-1$ και $2 \leq j \leq n$. Αυτό μας δίνει ένα σύστημα $j-1$ εξισώσεων με $j-1$ αγνώστους $p_{1,j}, \dots, p_{j-1,j}$ με πίνακα συντελεστών τον

$$\begin{pmatrix} a_{11} & \cdots & a_{1,j-1} \\ \vdots & & \vdots \\ a_{j-1,1} & \cdots & a_{j-1,j-1} \end{pmatrix}$$

ο οποίος έχει ορίζουσα $\Delta_{j-1}(A)$. Αν όλα τα $\Delta_{j-1}(A) \neq 0$ τότε τα παραπάνω συστήματα έχουν μοναδικές λύσεις από τις οποίες υπολογίζεται με μοναδικό τρόπο ο πίνακας P . Δηλαδή αν ισχύει το 2. τότε ισχύει το 1.

Αντιστρόφως, αν ισχύει το 1. και $a^1, \dots, a^n$ είναι οι στήλες του πίνακα A και $b^1, \dots, b^n$ είναι οι στήλες του πίνακα B . Έχουμε δείξει ότι ισχύει

$$b^1 = a^1$$

$$b^r = \sum_{j=1}^{r-1} p_{jr} a^j + a^r \quad \text{για } r > 1$$

Για σταθερό k , $1 \leq k \leq n$ έχουμε ότι η r -στήλη του πίνακα

$$\begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk} \end{pmatrix}$$

δίνεται προσθέτοντας στην r στήλη του πίνακα

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

ένα γραμμικό συνδυασμό των άλλων στηλών, ένας μετασχηματισμός που διατηρεί την ορίζουσα. Αυτό, μαζί με το γεγονός ότι ο πίνακας B είναι τριγωνικός δίνει ότι

$$\Delta_k(A) = \Delta_k(B) = b_{11}b_{22}\cdots b_{kk} \text{ για κάθε } 1 \leq k \leq n.$$

Αφού ο A και ο P είναι αντιστρέψιμοι και ο B είναι αντιστρέψιμος οπότε έχουμε

$$\det(B) = b_{11}b_{22}\cdots b_{nn} \neq 0$$

άρα και όλοι οι $\Delta_k(B) = \Delta_k(A) \neq 0$. □

Θεώρημα VIII.2.20. Μια μιάμιση μορφή f σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V με πίνακα $A = (a_{ij})$, $a_{ij} = f(v_j, v_i)$ ως προς την διατεταγμένη βάση $B = \{v_1, \dots, v_n\}$ είναι θετική αν και μόνο αν $A = A^*$ και όλες οι ελάχιστες ορίζουσες $\Delta_k(A) > 0$.

Απόδειξη. Υποθέτουμε ότι $A = A^*$ και ότι $\Delta_k(A) > 0$ για κάθε $1 \leq k \leq n$. Σύμφωνα με το λήμμα [VIII.2.19](#) υπάρχει μοναδικός άνω τριγωνικός πίνακας P με μονάδες στην διαγώνιο ώστε ο $B = AP$ να είναι κάτω τριγωνικός. Ο πίνακας P^* είναι κάτω τριγωνικός συνεπώς ο πίνακας $P^*B = P^*AP$ είναι επίσης κάτω τριγωνικός. Αφού ο A^* είναι αυτοσυζυγής και ο πίνακας $D = P^*AP$ είναι αυτοσυζυγής και είναι και τριγωνικός, συνεπώς είναι διαγώνιος σύμφωνα με την πρόταση [VII.9.10](#). Έχουμε

$$\Delta_k(D) = \Delta_k(P^*A) = \Delta_k(B) = \Delta_k(A),$$

και ο πίνακας $D = \text{diag}(d_{11}, \dots, d_{nn})$ έχει $\Delta_k(D) = d_{11}d_{22}\cdots d_{kk}$. Η σχέση $\Delta_k(D) > 0$ για $1 \leq k \leq n$ μας δίνει διαδοχικά ότι $d_{kk} > 0$ για κάθε k .

Έχουμε τώρα ότι για τον πίνακα A της f ως προς την βάση B ότι $D = P^*AP$ είναι ο πίνακας της f ως προς την βάση $B' = \{v'_1, \dots, v'_n\}$ η οποία ορίζεται από

$$v'_k = \sum_{i=1}^n p_{ik} v_i$$

και αφού ο D είναι διαγώνιος με θετικά στοιχεία στην διαγώνιο έχουμε ότι

$$X^*DX > 0 \text{ για κάθε } X \neq 0.$$

Αντιστρόφως, αν f είναι θετική έχουμε ότι $A = A^*$. Θεωρούμε τον υπόχωρο V_k που παράγεται από τα k πρώτα διανύσματα $B_k := \{v_1, \dots, v_k\}$ της βάσης B και f_k τον περιορισμό της f στον $V_k \times V_k$. Είναι σαφές ότι η f_k είναι θετική μορφή στο V_k και στην βάση B_k αναπαρίσταται από τον πίνακα

$$A_k := \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}.$$

Σύμφωνα με την πρόταση [VIII.2.17](#) υπάρχει αντιστρέψιμος πίνακας P_k , ώστε $A_k = P_k^*P_k$. Τότε όμως έχουμε

$$\Delta_k(A) = \det(A_k) = \det(P_k^*P_k) = \det(P_k^*)\det(P_k) = \overline{\det(P_k)}\det(P_k) = |\det(P_k)|^2 > 0.$$

□

Έστω V πεπερασμένης διάστασης διανυσματικός χώρος με εσωτερικό γινόμενο. Αν f είναι μια μη-αρνητική μιάμιση μορφή, υπάρχει μοναδική αυτοσυζυγής γραμμική συνάρτηση $L_f : V \rightarrow V$, ώστε

$$f(v, w) = \langle L_f(v), w \rangle \text{ για κάθε } v, w \in V,$$

ώστε $\langle T_f(v), v \rangle \geq 0$.

Ορισμός VIII.2.21. Μία γραμμική συνάρτηση $L : V \rightarrow V$, στον πεπερασμένης διάστασης διανυσματικό χώρο V με εσωτερικό γινόμενο θα λέγεται μη-αρνητική αν $L = L^*$ και $\langle Lv, v \rangle \geq 0$ για κάθε $v \in V$. Θα λέγεται θετική αν είναι μη-αρνητική και επιπρόσθετα $\langle Lv, v \rangle = 0$ αν και μόνο αν $v = 0$.

Ασκήσεις

Άσκηση VIII.1 Δείξτε ότι για δύο μιάμιση μορφές f, g η εξίσωση

$$\langle f, g \rangle = \text{tr}(L_g^* L_f)$$

ορίζει ένα εσωτερικό γινόμενο στον διανυσματικό χώρο των μιάμιση μορφών.

Άσκηση VIII.2 Έστω f μια μιάμιση μορφή σε ένα πεπερασμένης διάστασης διανυσματικό χώρο με εσωτερικό γινόμενο. Δείξτε ότι υπάρχει μια ορθοκανονική βάση του V στην οποία ο πίνακας της f είναι άνω τριγωνικός.

Άσκηση VIII.3 Δείξτε ότι μια μιάμιση μορφή είναι ερμητιανή αν και μόνο αν $f(v, v) \in \mathbb{R}$ για κάθε $v \in V$.

VIII.2.2 Διγραμμικές μορφές

Θα μελετήσουμε τώρα την θεωρία των διγραμμικών μορφών όπως αυτές ορίστηκαν στο [VIII.2.1](#). Είναι σαφές ότι στην περίπτωση ενός πραγματικού διανυσματικού χώρου οι μιάμιση μορφές ταυτίζονται με τις διγραμμικές. Ας δούμε μερικά ακόμα παραδείγματα διγραμμικών μορφών που δουλεύουν εξίσου καλά και στην περίπτωση οποιουδήποτε σώματος $\mathbb{F}$.

Παρατήρηση VIII.2.22. Παρατηρούμε ότι το σύνολο $L(V, V, \mathbb{F})$ των διγραμμικών μορφών $V \times V \rightarrow \mathbb{F}$, όπου V διανυσματικός χώρος πάνω από το σώμα $\mathbb{F}$ είναι διανυσματικός χώρος.

Παράδειγμα VIII.2.23. Έστω ένας διανυσματικός χώρος V και έστω $L_1, L_2 : V \rightarrow \mathbb{F}$. Η συνάρτηση $V \times V \rightarrow \mathbb{F}$ $f(v, w) = L_1(v)L_2(w)$ είναι διγραμμική.

Παράδειγμα VIII.2.24. Θεωρούμε τον διανυσματικό χώρο $V = \mathbb{F}^{n,m}$ των $n \times m$ πινάκων πάνω από το σώμα $\mathbb{F}$. Για ένα σταθερό πίνακα $A \in \mathbb{F}^{m,m}$ η συνάρτηση

$$\begin{aligned} V \times V &\longrightarrow \mathbb{F} \\ (X, Y) &\longmapsto f_A(X, Y) = \text{tr}(X^t A Y) \end{aligned}$$

είναι μια διγραμμική μορφή στον χώρο V .

Μπορούμε στην πραγματικότητα να περιγράψουμε όλες τις τετραγωνικές μορφές σε ένα διανυσματικό χώρο ως εξής: Θεωρούμε μια διατεταγμένη βάση $B = \{v_1, \dots, v_n\}$ του V και $f(v, w)$ μια διγραμμική μορφή. Αν $v = x_1 v_1 + \dots + x_n v_n$, $w = y_1 v_1 + \dots + y_n v_n$ τότε έχουμε

$$f(v, w) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j f(v_i, v_j).$$

Οπότε θέτοντας $A = (a_{ij})$ με $a_{ij} = f(v_i, v_j)$ έχουμε ότι

$$f(v, w) = [v]_B^t A [w]_B.$$

Αντιστρόφως, κάθε πίνακας A ορίζει μια διγραμμική μορφή με τον παραπάνω τρόπο.

Ορισμός VIII.2.25. Ο πίνακας A θα λέγεται ο πίνακας της διγραμμικής μορφής f ως προς την διατεταγμένη βάση B και θα τον συμβολίζουμε με $[f]_B$.

Θεώρημα VIII.2.26. Για ένα πεπερασμένης διάστασης διανυσματικό χώρο V και μια διατεταγμένη βάση B , η συνάρτηση

$$\begin{aligned} L(V, V, \mathbb{F}) &\longrightarrow \mathbb{F}^{n,n} \\ f &\longmapsto [f]_B \end{aligned}$$

είναι ισομορφισμός διανυσματικών χώρων.

Απόδειξη. Δεδομένης μιας διγραμμικής μορφής f ο πίνακας της $[f]_B$ είναι μονοσήμαντα ορισμένος, ενώ για κάθε πίνακα $A \in \mathbb{F}^{n,n}$ υπάρχει μοναδική διγραμμική συνάρτηση που αντιστοιχεί σε αυτόν. Τέλος, η συνάρτηση αντιστοίχισης είναι γραμμική δηλαδή

$$[\lambda_1 f_1 + \lambda_2 f_2]_B = \lambda_1 [f_1]_B + \lambda_2 [f_2]_B \text{ για κάθε } \lambda_1, \lambda_2 \in \mathbb{F}, f_1, f_2 \in L(V, V, \mathbb{F}).$$

□

Πόρισμα VIII.2.27. Αν $B = \{v_1, \dots, v_n\}$ είναι μια διατεταγμένη βάση του V και $B^* = \{v_1^*, \dots, v_n^*\}$ είναι η δυική βάση του $V^* = \text{Hom}(V, \mathbb{F})$, δηλαδή $v_i^*(v_j) = \delta_{ij}$ τότε οι n^2 το πλήθος διγραμμικές μορφές

$$f_{ij}(v, w) = v_i^*(v) v_j^*(w), 1 \leq i, j \leq n$$

αποτελούν μια βάση του χώρου $L(V, V, \mathbb{F})$. Η διάσταση του χώρου $L(V, V, \mathbb{F})$ είναι n^2 .

Απόδειξη. Παρατηρούμε ότι οι f_{ij} είναι πράγματι διγραμμικές μορφές, όπως αυτές ορίστηκαν στο παράδειγμα **VIII.2.23** και μάλιστα

$$f_{ij}(v, w) = x_i y_j$$

όπου $[v]_B = (x_1, \dots, x_n)^t$, $[w]_B = (y_1, \dots, y_n)^t$. Η γενική διγραμμική μορφή που αντιστοιχεί στον πίνακα $A = (a_{ij})$ γράφεται ως

$$f_A(v, w) = [v]_B^t A [w]_B = \sum_{i=1}^n \sum_{j=1}^n a_{ij} f_{ij}.$$

Δηλαδή, οι f_{ij} είναι n^2 το πλήθος μορφές, γραμμικά ανεξάρτητες που παράγουν τον χώρο $L(V, V, \mathbb{F})$. □

Παρατήρηση VIII.2.28. Αν P είναι ο πίνακας αλλαγής βάσης ώστε $[v]_B = P[v]_{B'}$, από την διατεταγμένη βάση B στην διατεταγμένη βάση B' τότε

$$f(v, w) = [v]_B^t [f]_B [w]_B = (P[v]_{B'})^t [f]_B P[v]_{B'}$$

και από το μονοσήμαντο του πίνακα της διγραμμικής συνάρτησης έχουμε ότι

$$[f]_{B'} = P^t [f]_B P.$$

Παρατήρηση VIII.2.29. Αν σταθεροποιήσουμε ένα $v \in V$ τότε η διγραμμική μορφή f δίνει μια συνάρτηση

$$\begin{aligned} L_f(v) : V &\longrightarrow \mathbb{F} \\ w &\longmapsto f(v, w) \end{aligned}$$

δηλαδή δίνει μια συνάρτηση

$$\begin{aligned} V &\longrightarrow V^* \\ v &\longmapsto L_f(\cdot) = f(v, \cdot) \end{aligned}$$

Ομοίως, αν σταθεροποιήσουμε ένα $w \in V$ έχουμε

$$\begin{aligned} R_f(w) : V &\longrightarrow \mathbb{F} \\ v &\longmapsto f(v, w) \end{aligned}$$

δηλαδή δίνει μια συνάρτηση

$$\begin{aligned} V &\longrightarrow V^* \\ w &\longmapsto R_f(\cdot) = f(\cdot, w) \end{aligned}$$

Θεώρημα VIII.2.30. Έστω μία διγραμμική μορφή σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V και έστω $L_f, R_f : V \rightarrow V^*$. Έχουμε ότι $r(L_f) = r(R_f)$.

Απόδειξη. Αρκεί να αποδείξουμε ότι $\dim \ker R_f = \dim \ker L_f$. Γράφουμε την $f(v, w)$ ως

$$f(v, w) = [v]_B^t A [w]_B.$$

Αν $w \in \ker R_f$ τότε έχουμε ότι $f(v, w) = 0$ για κάθε $v \in V$, δηλαδή $A[w]_B = \mathbf{0} \in \mathbb{F}^{n,n}$. Με τον ίδιο τρόπο αν $v \in \ker L_f$ έχουμε ότι $f(v, w) = 0$ για κάθε $w \in V$, δηλαδή $A^t[v]_B = \mathbf{0}$. Όμως $\dim \ker A = \dim \ker A^t$ και το ζητούμενο έπεται. $\square$

Ορισμός VIII.2.31. Η τάξη μιας διγραμμικής μορφής $r(f)$ ορίζεται να είναι ο $r(L_f) = r(R_f) = r([f]_B)$.

Πόρισμα VIII.2.32. Για μια διγραμμική μορφή $f : V \times V \rightarrow \mathbb{F}$ τα παρακάτω είναι ισοδύναμα

1. $r(f) = \dim V$.
2. Για κάθε $v \neq 0, v \in V$ υπάρχει μη-μηδενικό $w \in V$ ώστε $f(v, w) \neq 0$.
3. Για κάθε $w \neq 0, w \in V$ υπάρχει μη-μηδενικό $v \in V$ ώστε $f(v, w) \neq 0$.

Απόδειξη. Το 2. λέει ότι $\ker L_f = \{0\}$ ενώ το 3. ότι $\ker R_f = \{0\}$. Οι γραμμικές συναρτήσεις L_f και R_f έχουν μηδενικό πυρήνα αν και μόνο αν η τάξη τους, άρα και η τάξη της f είναι n . $\square$

Ορισμός VIII.2.33. Μια διγραμμική μορφή θα λέγεται μη-ιδιόμορφη αν ισχύει μία από τις ισοδύναμες προτάσεις του πορίσματος **VIII.2.32**.

Ορισμός VIII.2.34. Μια διγραμμική μορφή θα λέγεται συμμετρική αν μόνο αν

$$f(v, w) = f(w, v) \text{ για κάθε } v, w \in V.$$

Παρατήρηση VIII.2.35. Παρατηρούμε ότι η $f(v, w) = [v]_B^t [f]_B [w]_B$ είναι συμμετρική αν και μόνο αν $[f]_B = [f]_B^t$.

Σε κάθε συμμετρική διγραμμική μορφή αντιστοιχεί μια τετραγωνική μορφή

$$q(v) = f(v, v) = [v]_B^t [f]_B [v]_B. \quad (\text{VIII.12})$$

Η τετραγωνική μορφή καθορίζει την διγραμμική μορφή σύμφωνα με την ταυτότητα

$$f(v, w) = \frac{1}{4}q(v+w) - \frac{1}{4}q(v-w).$$

Θεώρημα VIII.2.36. Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο πάνω από το σώμα $\mathbb{F}$, χαρακτηριστικής 0 υπάρχει μια διατεταγμένη βάση του V ώστε ο πίνακας $[f]_B$ της συμμετρικής διγραμμικής μορφής f να είναι διαγώνιος.

Απόδειξη. Ψάχνουμε να βρούμε διατεταγμένη βάση $B = \{v_1, \dots, v_n\}$ ώστε $f(v_i, v_j) = 0$ αν $i \neq j$. Αν $f = 0$ ή $n = \dim V = 1$ τότε το θεώρημα είναι προφανές. Αν $q(v) = f(v, v) = 0$ για κάθε $v \in V$ τότε η σχέση **(VIII.12)** μας δίνει ότι $f = 0$. Υποθέτουμε ότι υπάρχει $v \in V$ ώστε $q(v) \neq 0$. Θεωρούμε τον χώρο $W = \langle v \rangle$ και ορίζουμε τον χώρο $W^\perp = \{w \in V : f(v, w) = 0\}$. Θα δείξουμε ότι $V = W \oplus W^\perp$. Είναι σαφές ότι $W \cap W^\perp = \{0\}$. Πράγματι, για ένα στοιχείο στον W της μορφής λv έχουμε ότι $0 = f(\lambda v, \lambda v) = \lambda^2 f(v, v)$ και αφού $f(v, v) \neq 0$ έχουμε ότι $\lambda = 0$. Κάθε διάνυσμα στο V είναι άθροισμα ενός στοιχείου του W και ενός στοιχείου του $W^\perp$. Πράγματι, για $w \in V$, θέτουμε

$$w_1 = w - \frac{f(w, v)}{f(v, v)}v$$

και παρατηρούμε ότι

$$f(w_1, v) = f(w, v) - \frac{f(w, v)}{f(v, v)}f(v, v) = 0.$$

Συνεπώς $w_1 \in W^\perp$ και $w = w_1 + \frac{f(w, v)}{f(v, v)}v$. Ο περιορισμός της f στο $W^\perp$ είναι μια συμμετρική διγραμμική μορφή και αφού $\dim W^\perp = n - 1$ το αποτέλεσμα προκύπτει επαγωγικά. Μπορούμε δηλαδή να υποθέσουμε ότι ο $W^\perp$ έχει μια διατεταγμένη βάση $v_1, \dots, v_{n-1}$, ώστε $f(v_i, v_j) = 0$ για $i \neq j$, $1 \leq i, j \leq n - 1$. Η ζητούμενη βάση είναι η $\{v_1, \dots, v_{n-1}, v_n\}$. $\square$

Πόρισμα VIII.2.37. Αν $\mathbb{F}$ είναι ένα υπόσωμα του $\mathbb{C}$ και $A \in \mathbb{F}^{n,n}$, $A^t = A$, τότε υπάρχει αντιστρέψιμος πίνακας P , ώστε $P^t A P$ να είναι διαγώνιος.

Παρατήρηση VIII.2.38. Στην περίπτωση που $\mathbb{F} = \mathbb{R}$ τον πίνακα P μπορούμε να τον διαλέξουμε ώστε $P^t = P^{-1}$.

Θεώρημα VIII.2.39. Αν V είναι διανυσματικός χώρος πάνω από το $\mathbb{C}$ και f είναι μια συμμετρική διγραμμική μορφή τότε υπάρχει διατεταγμένη βάση $B = \{v_1, \dots, v_n\}$ ώστε $f(v_i, v_j) = 0$ αν $i \neq j$,

$$f(v_i, v_i) = \begin{cases} 1 & \text{αν } 1 \leq i \leq r \\ 0 & \text{αν } i > r \end{cases}$$

Απόδειξη. Υπάρχει μια βάση $B = \{v_1, \dots, v_n\}$ στην οποία ο πίνακας της f να είναι διαγώνιος. Αν ο $[f]_B$ έχει τάξη $0 \leq r \leq n$ τότε στην διαγώνιο έχουμε ακριβώς r το πλήθος μη-μηδενικά στοιχεία. Αναδιατάσσοντας την βάση μπορούμε να υποθέσουμε ότι αυτά είναι τα r πρώτα. Στην συνέχεια θέτουμε

$$v'_j = \frac{1}{\sqrt{f(v_i, v_i)}} v_i \text{ για κάθε } 1 \leq j \leq r.$$

Η ζητούμενη βάση είναι η $v'_1, \dots, v'_r, v_{r+1}, \dots, v_n$. □

Στους πραγματικούς αριθμούς οι αρνητικοί αριθμοί δεν έχουν τετραγωνικές ρίζες. Μπορούμε όμως να αποδείξουμε ότι

Θεώρημα VIII.2.40. Έστω V ένας πραγματικός διανυσματικός χώρος διάστασης n και έστω f διγραμμική συμμετρική μορφή. Υπάρχει μια βάση B , ώστε ο πίνακας $[f]_B$ να είναι διαγώνιος και ώστε

$$f(v_i, v_i) = \begin{cases} \pm 1 & \text{αν } 1 \leq i \leq r \\ 0 & \text{αν } i > r \end{cases}$$

Επιπλέον το πλήθος των διανυσμάτων της βάσης ώστε $f(v_i, v_i) = 1$ είναι ανεξάρτητο της επιλογής της βάσης.

Απόδειξη. Όπως και πριν υπάρχει μια βάση $B = \{v_1, \dots, v_n\}$ ώστε

$$\begin{aligned} f(v_i, v_j) &= 0 \text{ αν } i \neq j \\ f(v_i, v_i) &\neq 0 \text{ αν } 1 \leq i \leq r \\ f(v_i, v_i) &= 0 \text{ αν } i > r \end{aligned}$$

Θέτουμε

$$v'_i = |f(v_i, v_i)|^{-1/2} v_i \text{ αν } 1 \leq i \leq r.$$

Η ζητούμενη βάση είναι η $\{v'_1, \dots, v'_r, v_{r+1}, \dots, v_n\}$.

Έστω p το πλήθος των $v'_1, \dots, v'_r$ ώστε $f(v'_i, v'_i) = 1$. Θα δείξουμε ότι το p είναι ανεξάρτητο από την επιλογή της βάσης που κάναμε αν έχουμε την ζητούμενη μορφή. Θέτουμε V^+ τον χώρο που παράγεται από τα διανύσματα $v'_1, \dots, v'_r$ και $f(v'_i, v'_i) = 1$ και V^- τον χώρο που παράγεται από τα $v'_1, \dots, v'_r$ ώστε $f(v'_i, v'_i) = -1$. Ισχύει ότι $p = \dim V^+$. Έστω V_0 ο υπόχωρος που παράγεται από τα $v_{r+1}, \dots, v_n$.

Για κάθε μη-μηδενικό διάνυσμα $v \in V^+$ έχουμε $f(v, v) > 0$ ενώ για κάθε μη-μηδενικό διάνυσμα $v \in V^-$ έχουμε $f(v, v) < 0$. Ισχύει ότι

$$V = V^+ \oplus V^- \oplus V^0. \quad (\text{VIII.13})$$

Έστω ένας υπόχωρος W του V ώστε $f(v, v) > 0$ για κάθε $v \in W$. Ισχύει ότι αν έχουμε ένα άθροισμα

$$w + v^- + v_0 = 0, w \in W, v^- \in V^-, v_0 \in V^0,$$

τότε $w = v^- = v_0 = 0$. Πράγματι,

$$\begin{aligned} 0 &= f(w, w + v^- + v_0) = f(w, w) + f(w, v^-) + f(w, v_0) \\ 0 &= f(v^-, w + v^- + v_0) = f(v^-, w) + f(v^-, v^-) + f(v^-, v_0). \end{aligned}$$

Παρατηρούμε ότι αν $v_0 \in V^0$ τότε $f(v, v_0) = 0$ για κάθε $v \in V$. Συνεπώς οι παραπάνω σχέσεις γίνονται

$$\begin{aligned} 0 &= f(w, w) + f(w, v^-) \\ 0 &= f(v^-, w) + f(v^-, v^-). \end{aligned}$$

Επειδή η διγραμμική είναι συμμετρική η παραπάνω σχέση δίνει ότι $f(w, w) = f(v^-, v^-)$ και επειδή $f(w, w) \geq 0$ και $f(v^-, v^-) \leq 0$ έχουμε ότι $f(w, w) = f(v^-, v^-) = 0$ και συνεπώς $w = v^- = 0$ το οποίο δίνει και ότι $v_0 = 0$. Η σχέση (VIII.13) μαζί με την ανεξαρτησία των W, V^-, V^0 επιβάλλει ότι $\dim W \leq \dim V^+$. Άρα αν σχηματίσουμε τους ίδιους χώρους V_1^+, V_1^-, V_1^0 ως προς μια άλλη βάση θα έχουμε ότι $\dim V_1^+ \leq \dim V^+$ και για συμμετρικούς λόγους καταλήγουμε στο ότι $\dim V^+ = \dim V_1^+$. $\square$

Παρατήρηση VIII.2.41.

$$\dim V^+ + \dim V^- = r(f).$$

Η ποσότητα $\dim V^+ - \dim V^-$ λέγεται υπογραφή της διγραμμικής μορφής και μαζί με το n και $r(f)$ μπορούμε να ανακτήσουμε τις διαστάσεις των χώρων V^+, V^-, V^0 .

VIII.3 Φασματική Θεωρία

Θεώρημα VIII.3.1. Έστω $L : V \rightarrow V$ μία κανονική γραμμική συνάρτηση, όπου V είναι πεπερασμένης διάστασης μιγαδικός διανυσματικός χώρος ή είναι μια αυτοσυζυγής γραμμική συνάρτηση σε ένα πεπερασμένης διάστασης πραγματικό χώρο. Θεωρούμε το σύνολο των διαφορετικών ανά δύο ιδιοτιμών της L , $\lambda_1, \dots, \lambda_k$, $k \leq \dim V$. Έστω E_i ο ιδιόχωρος της ιδιοτιμής λ_i . Τότε οι ιδιόχωροι E_i είναι ανά δύο κάθετοι. Επίσης, αν $\pi_i : V \rightarrow E_i$ είναι η προβολή του V στον χώρο E_i , τότε

$$L = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k.$$

Απόδειξη. Στην περίπτωση που η L είναι κανονική απεικόνιση επί μιγαδικού διανυσματικού χώρου για να δείξουμε ότι οι χώροι E_i, E_j είναι κάθετοι αν $i \neq j$, διαλέγουμε $v_i \in E_i, v_j \in E_j$ και υπολογίζουμε

$$\lambda_i \langle e_i, e_j \rangle = \langle Le_i, e_j \rangle = \langle e_i, L^* e_j \rangle \stackrel{\text{VII.9.8}}{=} \langle e_i, \bar{\lambda}_j e_j \rangle = \lambda_j \langle e_i, e_j \rangle$$

από όπου προκύπτει το ζητούμενο, αφού $\lambda_i \neq \lambda_j$.

Στην περίπτωση που η L είναι αυτοσυζυγής έχουμε δείξει στην πρόταση VII.9.4 με παρόμοιο τρόπο ότι οι ιδιόχωροι είναι κάθετοι μεταξύ τους.

Αφού οι κανονικές γραμμικές συναρτήσεις είναι διαγωνοποιήσιμες ο χώρος V έχει μια βάση αποτελούμενη από ιδιοδιανύσματα οπότε

$$V = E_1 + \dots + E_k.$$

Έχουμε ότι αν $v_i \in E_i$ και $v_1 + \dots + v_k = 0$, τότε

$$0 = \langle v_i, v_1 + \dots + v_k \rangle = \langle v_i, v_i \rangle = \|v_i\|^2,$$

άρα κάθε $v_i = 0$ και το άθροισμα είναι ευθύ. Συνεπώς

$$\text{Id}_V = \pi_1 + \dots + \pi_k$$

από όπου έχουμε

$$L = L \circ \pi_1 + \dots + L \circ \pi_k = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k.$$

□

Πρόταση VIII.3.2. Οι προβολές στους ιδιόχωρους είναι πολυωνυμικές συναρτήσεις του L . Για

$$p_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x - \lambda_j}{\lambda_j - \lambda_i} \right)$$

έχουμε $p_i(L) = \pi_i$.

Απόδειξη. Αφού $\pi_i \pi_j = \delta_{ij}$ έχουμε ότι

$$L^2 = \lambda_1^2 \pi_1 + \dots + \lambda_k^2 \pi_k,$$

από όπου επαγωγικά προκύπτει ότι

$$L^n = \lambda_1^n \pi_1 + \dots + \lambda_k^n \pi_k \text{ για κάθε } n \geq 0.$$

Συνεπώς, για ένα οποιοδήποτε πολυώνυμο

$$f(x) = \sum_{n=0}^d a_n x^n,$$

έχουμε ότι

$$f(L) = \sum_{n=0}^d a_n L^n = \sum_{n=0}^d a_n \sum_{j=1}^k \lambda_j^n \pi_j = \sum_{j=1}^k \left(\sum_{n=0}^d a_n \lambda_j^n \right) \pi_j = \sum_{j=1}^k f(\lambda_j) \pi_j.$$

Αφού $p_i(\lambda_j) = \delta_{ij}$, έχουμε ότι $p_i(L) = \pi_i$.

□

Παρατήρηση VIII.3.3. Οι προβολές $\pi_1, \dots, \pi_k$ εξαρτώνται από τον L . Την γραφή

$$\text{Id}_V = \pi_1 + \dots + \pi_k$$

θα την λέμε ανάλυση της ταυτότητας ορισμένη από τον L .

Ορισμός VIII.3.4. Θεωρούμε μια διαγωνοποιήσιμη κανονική γραμμική συνάρτηση $L : V \rightarrow V$ σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V με εσωτερικό γινόμενο και έστω

$$L = \sum_{j=1}^k \lambda_j \pi_j$$

η φασματική της ανάλυση. Αν $f : \mathbb{C} \supseteq U \rightarrow \mathbb{C}$ είναι μια συνάρτηση που το πεδίο ορισμού της περιέχει τα $\{\lambda_1, \dots, \lambda_k\}$, τότε μπορούμε να ορίσουμε την γραμμική συνάρτηση $f(L) : V \rightarrow V$

ως εξής:

$$f(L) = \sum_{j=1}^k f(\lambda_j) \pi_j.$$

Θεώρημα VIII.3.5. Έστω $L : V \rightarrow V$ διαγωνοποιήσιμη κανονική γραμμική συνάρτηση με σύνολο ιδιοτιμών $S = \{\lambda_1, \dots, \lambda_k\}$, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος με εσωτερικό γινόμενο. Αν $f : \mathbb{C} \supset U \rightarrow \mathbb{C}$ συνάρτηση με το πεδίο ορισμού της U να περιέχει το S . Τότε ο $f(L)$ είναι διαγωνοποιήσιμη κανονική γραμμική συνάρτηση με σύνολο ιδιοτιμών $f(S)$. Αν $Q : V \rightarrow V'$ είναι μοναδιαία γραμμική συνάρτηση και $L' = QLQ^{-1}$, τότε το σύνολο ιδιοτιμών της L' είναι επίσης το S και

$$f(L') = Qf(L)Q^{-1}.$$

Απόδειξη. Για να αποδείξουμε ότι η $f(L)$ είναι κανονική υπολογίζουμε ότι

$$f(L)^* = \sum_{j=1}^k \overline{f(\lambda_j)} \pi_j,$$

συνεπώς

$$f(L)f(L)^* = \sum_{j=1}^k \sum_{i=1}^k f(\lambda_i) \overline{f(\lambda_j)} \pi_j \pi_i \stackrel{\pi_j \pi_i = \delta_{ij} \pi_i}{=} \sum_{i=1}^k |f(\lambda_i)|^2 \pi_i = f(L)^* f(L).$$

Επιπλέον είναι σαφές ότι για κάθε $v \in \pi_j(V)$ έχουμε ότι

$$f(L)v = f(\lambda_j)v,$$

συνεπώς το σύνολο $f(S)$ περιέχεται στο σύνολο των ιδιοτιμών του $f(L)$. Αντιστρόφως, αν $v \neq 0$ και

$$f(L)v = \lambda v,$$

τότε αφού $v = \sum_{j=1}^k \pi_j v$ έχουμε ότι

$$f(L)v = \sum_{j=1}^k f(L) \pi_j v = \sum_{j=1}^k f(\lambda_j) \pi_j v = \sum_{j=1}^k \lambda \pi_j v$$

Από εδώ έχουμε ότι

$$\left\| \sum_{j=1}^k f(\lambda_j) - \lambda \pi_j v \right\|^2 = \sum_{j=1}^k |f(\lambda_j) - \lambda|^2 \|\pi_j v\|^2 = 0$$

δηλαδή $f(\lambda_j) - \lambda = 0$ ή $\pi_j v = 0$. Συνεπώς $f(\lambda_i) = \lambda$ για κάποιο i και συνεπώς το $f(S)$ είναι το σύνολο των ιδιοτιμών του $f(L)$.

Υποθέτουμε $Q : V \rightarrow V'$ είναι μοναδιαία γραμμική συνάρτηση και ότι $L' = QLQ^{-1}$. Τότε η εξίσωση

$$Lv = \lambda v$$

ισχύει αν και μόνο αν

$$L'Qv = \lambda Qv.$$

Συνεπώς, το S είναι επίσης το σύνολο των ιδιοτιμών του L' και η Q μεταφέρει τον ιδιόχωρο της ιδιοτιμής λ για τον L στον ιδιόχωρο της ιδιοτιμής λ για τον L' . Έτσι γράφουμε την φασματική ανάλυση για τον L

$$L' = \sum_{j=1}^j \lambda_j \pi'_j, \text{ όπου } \pi'_j = Q \pi_j Q^{-1}$$

και έχουμε

$$f(L') = \sum_{j=1}^k f(\lambda_j) \pi_j' = \sum_{j=1}^k f(\lambda_j) Q \pi_j Q^{-1} = Q \left(\sum_{j=1}^k f(\lambda_j) \pi_j \right) Q^{-1} = Q f(L) Q^{-1}.$$

□

Ασκήσεις

Άσκηση VIII.4 Έστω L μια διαγωνοποιήσιμη γραμμική συνάρτηση $L : V \rightarrow V$ σε ένα πεπερασμένης διάστασης χώρο V με εσωτερικό γινόμενο. Δείξτε ότι

1. L είναι αυτοσυζυγής αν και μόνο αν οι ιδιοτιμές της είναι πραγματικές.
2. L είναι μη-αρνητική αν και μόνο αν οι ιδιοτιμές της είναι μη-αρνητικές.
3. L είναι μοναδιαία αν και μόνο αν οι ιδιοτιμές της έχουν μέτρο 1.

IX.1 Αναπαραστάσεις ομάδων

Θα παρουσιάσουμε μια σύντομη εισαγωγή στην θεωρία αναπαραστάσεων πεπερασμένων ομάδων, ακολουθώντας το κλασικό βιβλίο του J. P. Serre [11] όπως και το βιβλίο του Γ. Αντωνιάδη [16] ο οποίος με εισήγαγε στην θεωρία αυτή.

Ορισμός IX.1.1. Μια αναπαράσταση της πεπερασμένης ομάδας G είναι ένας ομομορφισμός ομάδων

$$\rho : G \longrightarrow \mathrm{GL}(V)$$

όπου V είναι ένας k -διάστατος $\mathbb{F}$ -διανυσματικός χώρος. Η διάσταση k θα λέγεται βαθμός της παράστασης και θα συμβολίζεται με $\deg \rho$.

Αν θεωρήσουμε μία διατεταγμένη βάση του V μπορούμε να θεωρήσουμε το V ισόμορφο με το $\mathbb{F}^k$ και μπορούμε να περάσουμε σε μια αναπαράσταση δια πινάκων δηλαδή έναν ομομορφισμό ομάδων

$$\rho : G \longrightarrow \mathrm{GL}_k(\mathbb{F}).$$

Η επιλογή της βάσης καθορίζει την παραπάνω αναπαράσταση. Ο παρακάτω ορισμός κάνει τέτοιες αναπαραστάσεις ισοδύναμες.

Ορισμός IX.1.2. Δυο αναπαραστάσεις $\rho_1, \rho_2 : G \rightarrow \mathrm{GL}_k(\mathbb{F})$ θα λέγονται ισοδύναμες αν είναι ίδιου βαθμού και υπάρχει $Q \in \mathrm{GL}_k(\mathbb{F})$, ώστε

$$\rho_1(g) = Q^{-1} \rho_2(g) Q \text{ για κάθε } g \in G.$$

Παράδειγμα IX.1.3.

- Μια αναπαράσταση βαθμού 1 είναι ένας ομομορφισμός $\rho : G \rightarrow \mathbb{F}^*$, όπου $\mathbb{F}^* = \mathrm{GL}_1(\mathbb{F})$ είναι η πολλαπλασιαστική ομάδα του σώματος $\mathbb{F}$. Αφού κάθε στοιχείο g έχει πεπερασμένη τάξη, οι εικόνες $\rho(g)$ είναι ρίζες της μονάδας στο $\mathbb{F}$.
- Η ομαλή αναπαράσταση της G είναι μια αναπαράσταση $\rho : G \rightarrow \mathrm{GL}(V)$, στην οποία έχουμε ότι τα στοιχεία $e_g = \rho(g)e_1$ αποτελούν μια βάση του V , δηλαδή $\deg \rho = |G|$.

Ορισμός ΙΧ.1.4. Ένας υπόχωρος W του V θα λέγεται ρ -αναλλοίωτος αν $\rho(g)W \subset W$ για κάθε $g \in G$.

Μία αναπαράσταση $\rho : G \rightarrow GL(V)$ θα λέγεται ανάγωγη αν δεν υπάρχει ρ -αναλλοίωτος υπόχωρος W με $W \neq \{0\}, W \neq V$.

Ορισμός ΙΧ.1.5. Αν $\rho_i : G \rightarrow GL(W_i)$, $i = 1, 2$ αναπαραστάσεις της G με βαθμούς k_1 και k_2 τότε ορίζεται μια αναπαράσταση $\rho_1 \oplus \rho_2$

$$\rho_1 \oplus \rho_2 : G \rightarrow GL(W_1 \oplus W_2)$$

η οποία σε μορφή πινάκων γράφεται ως εξής:

$$\begin{aligned} \rho_1 \oplus \rho_2 : G &\longrightarrow GL_{k_1+k_2}(\mathbb{F}) \\ g &\longmapsto \left(\begin{array}{c|c} \rho_1(g) & \mathbf{0} \\ \hline \mathbf{0} & \rho_2(g) \end{array} \right) \end{aligned}$$

η οποία θα λέγεται το ευθύ άθροισμα των αναπαραστάσεων ρ_1, ρ_2 .

Θεώρημα ΙΧ.1.6. Έστω G πεπερασμένη ομάδα τάξης $|G|$ ώστε το $|G|$ να αντιστρέφεται στο $\mathbb{F}$. Αν $W < V$ είναι ένας ρ -αναλλοίωτος υπόχωρος τότε υπάρχει W_1 ο οποίος είναι ρ -αναλλοίωτος υπόχωρος του V με $W \oplus W_1 = V$, δηλαδή

$$\rho = \rho|_W \oplus \rho|_{W_1}.$$

Απόδειξη. Υπάρχει διανυσματικός χώρος W' ώστε $V = W \oplus W'$. Για παράδειγμα διαλέγουμε μια βάση $\{e_1, \dots, e_{\dim W}\}$ του W την οποία την επεκτείνουμε σε μια βάση $\{e_1, \dots, e_{\dim W}, e_1, \dots, e_n\}$ του V και θέτουμε $W' = \langle e_1, \dots, e_n \rangle$. Το παραπάνω ευθύ άθροισμα είναι ένα άθροισμα διανυσματικών χώρων, αλλά ο χώρος W' δεν είναι απαραίτητα ρ -αναλλοίωτος.

Θεωρούμε την προβολή

$$\begin{aligned} p_W : V &\longrightarrow W \\ v = w + w' &\longmapsto w \end{aligned}$$

Έχουμε ότι $\ker p = W'$ και $p|_W = \text{Id}_W$. Θεωρούμε την προβολή

$$p_1 = \frac{1}{|G|} \sum_{t \in G} \rho(t) p \rho(t^{-1})$$

για την οποία ισχύει ότι

- $p_1(v) \in W$ για κάθε $v \in V$. Πράγματι,

$$p_1(v) = \frac{1}{|G|} \sum_{t \in G} \rho(t) p \rho(t^{-1})(v) = \frac{1}{|G|} \sum_{t \in G} \rho(t) p v_1 = \frac{1}{|G|} \sum_{t \in G} \rho(t) w_1 \in W,$$

αφού ο W είναι ρ -αναλλοίωτος.

- $p_1|_W = \text{Id}_W$. Πράγματι, $\rho(t^{-1})w \in W$, συνεπώς $p \rho(t^{-1})w = \rho(t^{-1})w$ και έχουμε

$$p_1(w) = \frac{1}{|G|} \sum_{t \in G} \rho(t) p \rho(t^{-1})(w) = \frac{1}{|G|} \sum_{t \in G} \rho(t) \rho(t^{-1})(w) = w.$$

- Ο χώρος $W_1 = \ker \pi_1$ είναι ρ -αναλλοίωτος. Πράγματι, έστω $w_1 \in W_1$ και $g \in G$. Θα δείξουμε ότι $\rho(g)w_1 \in W_1$ δηλαδή ότι $\pi_1(\rho(g)w_1) = 0$. Έχουμε

$$\pi_1(\rho(g)w_1) = \frac{1}{|G|} \sum_{t \in G} \rho(t) \rho(t^{-1})(\rho(g)w_1) = \frac{1}{|G|} \sum_{t \in G} \rho(t) \rho(t^{-1}g)w_1$$

Θέτουμε $t_1 = g^{-1}t$ και η τελευταία σχέση δίνει

$$\pi_1(\rho(g)w_1) = \frac{1}{|G|} \sum_{t_1 \in G} \rho(gt_1) \rho(t_1^{-1})w_1 = \rho(g) \frac{1}{|G|} \sum_{t_1 \in G} \rho(t_1) \rho(t_1^{-1})w_1 = \rho(g)\pi_1(w_1) = 0,$$

αφού $w_1 \in \ker \pi_1$. Συνεπώς W_1 είναι ρ -αναλλοίωτος.

Το W_1 είναι ορθογώνιο συμπλήρωμα του W . Αν $v \in V$ τότε γράφουμε

$$v = \pi_1(v) + v - \pi_1(v).$$

Έχουμε ότι $\pi_1(v) \in W$ και $v - \pi_1(v) \in W_1$ αφού $\pi_1(v - \pi_1(v)) = \pi_1(v) - \pi_1^2(v) = \pi_1(v) - \pi_1(v) = 0$. Άρα $V = W + W_1$. Τέλος αν $v \in W \cap W_1$, από την μία έχουμε ότι $v = \pi_1(v) = 0$, άρα $W \cap W_1 = 0$ και το άθροισμα είναι ευθύ. $\square$

Πρόταση IX.1.7. Κάθε αναπαράσταση είναι ευθύ άθροισμα αναγωγών αναπαραστάσεων.

Απόδειξη. Επαγωγή στον βαθμό της αναπαράστασης και χρήση του θεωρήματος IX.1.6. $\square$

Ορισμός IX.1.8. Αν G είναι μια πεπερασμένη ομάδα και $\rho : G \rightarrow GL_k(\mathbb{F})$ (αντίστοιχα $\rho : G \rightarrow GL(V)$) είναι μια αναπαράσταση της. Ορίζουμε την συνάρτηση

$$\begin{aligned} \chi_\rho : G &\longrightarrow \mathbb{F} \\ g &\longmapsto \chi_\rho(g) = \text{tr}(\rho(g)) \end{aligned}$$

την οποία θα ονομάζουμε χαρακτήρα της g .

Παρατήρηση IX.1.9. 1. Ισοδύναμες αναπαραστάσεις έχουν τον ίδιο χαρακτήρα.

2. $\chi_\rho(e_G) = \deg \rho$
3. $\chi_\rho(g_1 g_2) = \chi_\rho(g_2 g_1)$ για κάθε $g_1, g_2 \in G$.
4. Αν $\deg \rho = 1$ τότε $\chi_\rho = \rho$.
5. $\chi_{\rho_1 + \rho_2} = \chi_{\rho_1} + \chi_{\rho_2}$.

Παράδειγμα IX.1.10. Ο χαρακτήρας r_G της ομαλής παράστασης δίνεται από

$$r_G(g) = \begin{cases} |G| & \text{αν } g = e_G \\ 0 & \text{αν } g \neq e_G. \end{cases}$$

Πράγματι, $r_G(e_G) = \deg G = |G|$. Αν $g \neq e_G$ έχουμε ότι στην βάση $\{e_t\}_{t \in G}$ έχουμε $ge_t = e_{gt}$ δηλαδή στον πίνακα της ομαλής παράστασης ως προς την βάση αυτή δεν υπάρχει κανένα στοιχείο στην διαγώνιο. Το αποτέλεσμα έπεται.

Πρόταση ΙΧ.1.11. Έστω $\rho_i : G \rightarrow GL_{k_i}(\mathbb{F})$, $i = 1, 2$ ανάγωγες αναπαραστάσεις της G . Έστω $A \in \mathbb{F}^{k_2, k_1}$ και υποθέτουμε ότι ισχύει

$$\rho_2(g)A = A\rho_1(g) \text{ για κάθε } g \in G.$$

Τότε A είναι ή αντιστρέψιμος πίνακας οπότε $k_1 = k_2$ και η ρ_1 είναι ισοδύναμη με την ρ_2 ή ο A είναι μηδενικός πίνακας.

Απόδειξη. Θεωρούμε τον $W_1 = \ker A$ και $W_2 = \operatorname{Im} A$. Ισχύει ότι

$$\rho_1(g)w \in W_1 \text{ για κάθε } w \in W_1 \text{ και } g \in G.$$

Πράγματι, για $w \in W_1$ έχουμε

$$A\rho_1(g)w = \rho_2(g)Aw = \rho_2(g)0 = 0$$

Ομοίως, για $w = Aw_1 \in W_2$ ισχύει ότι

$$\rho_2(g)w = \rho_2(g)Aw_1 = A\rho_1(g)w_1 \in W_2.$$

Αν δεν είναι μηδενικός πίνακας τότε $\ker A \neq \mathbb{F}^{k_1}$ συνεπώς $\ker(A) = 0$, διαφορετικά η ρ_1 δεν θα ήταν ανάγωγη. Επίσης $\operatorname{im}(A) \neq \{0\}$ άρα $\operatorname{im} A = \mathbb{F}^{k_2}$ επειδή η ρ_2 είναι ανάγωγη. $\square$

Πόρισμα ΙΧ.1.12. Έστω $\rho_i : G \rightarrow GL_{k_i}(\mathbb{F})$, $i = 1, 2$ δύο ανάγωγες μη-ισοδύναμες αναπαραστάσεις της G . Έστω $A \in \mathbb{F}^{k_2, k_1}$. Αν

$$A_0 = \sum_{t \in G} \rho_2(t^{-1})A\rho_1(t),$$

τότε $A_0 = 0$. Αν $\rho_2(t) = (a_{ij}(t))$ και $\rho_1(t) = (b_{ij}(t))$, τότε για κάθε $1 \leq i, j \leq k_2$ και $1 \leq \kappa, l \leq k_1$ έχουμε

$$\sum_{t \in G} a_{ij}(t)b_{lk}(t^{-1}) = 0. \quad (\text{IX.1})$$

Απόδειξη. Παρατηρούμε ότι για κάθε $g \in G$, έχουμε

$$\rho_2(g)A_0 = A_0\rho_1(g).$$

Πράγματι,

$$\rho_2(g)A_0 = \rho_2(g) \sum_{t \in G} \rho_2(t^{-1})A\rho_1(t) = \sum_{t \in G} \rho_2(gt^{-1})A\rho_1(t) \stackrel{t_1=gt^{-1}}{=} \sum_{t_1 \in G} \rho_2(t_1^{-1})A\rho_1(t_1)\rho_1(g) = A_0\rho_1(g).$$

Ο A_0 δεν μπορεί να είναι ισομορφισμός αφού οι ρ_1, ρ_2 δεν είναι ισοδύναμες, άρα $A_0 = 0$.

Διαλέγουμε $A = (a_{ij})$ με $a_{ij} = \delta_{j,\nu}\delta_{\mu,l}$ και υπολογίζουμε

$$\begin{aligned} 0 = c_{i\kappa} &= \sum_{t \in G} \sum_{\nu=1}^k \sum_{\mu=1}^k a_{i,\nu}(t^{-1})a_{\nu,\mu}b_{\mu,\kappa}(t) = \sum_{t \in G} \sum_{\nu=1}^k \sum_{\mu=1}^k a_{i,\nu}(t^{-1})\delta_{j,\nu}\delta_{\mu,l}b_{\mu,\kappa}(t) \\ &= \sum_{t \in G} a_{i,j}(t^{-1})b_{l,\kappa}(t) \end{aligned}$$

δηλαδή το ζητούμενο. $\square$

Θεώρημα IX.1.13. Αν $\mathbb{F}$ είναι αλγεβρικά κλειστό σώμα και $\rho : G \rightarrow GL_k(\mathbb{F})$ ανάγωγη αναπαράσταση. Αν $A \in \mathbb{F}^{k,k}$ ώστε $A\rho(g) = \rho(g)A$ για κάθε $g \in G$ τότε

$$\rho(g) = \lambda \cdot \mathbb{I}_k, \text{ με } \lambda \in \mathbb{F}.$$

Απόδειξη. Το χαρακτηριστικό πολυώνυμο $\text{ch}_A(x)$ του πίνακα A έχει μια ρίζα λ συνεπώς ο $A_1 = A - \lambda \mathbb{I}_n$ δεν είναι ισομορφισμός. Όμως ισχύει ότι $A_1 \rho(g) = \rho(g) A_1$ συνεπώς ο $A_1 = 0$, δηλαδή το ζητούμενο. $\square$

Στην συνέχεια υποθέτουμε ότι το σώμα $\mathbb{F}$ έχει χαρακτηριστική μηδέν και ότι είναι αλγεβρικά κλειστό.

Πόρισμα IX.1.14. Έστω G αβελιανή ομάδα. Κάθε ανάγωγη αναπαράσταση της G έχει βαθμό 1.

Απόδειξη. Το προηγούμενο θεώρημα για $g \in G$ σταθερό δίνει ότι $\rho(g) = \lambda_g \mathbb{I}_k$. Δηλαδή ο $\rho(g) = \text{diag}(\lambda_g, \dots, \lambda_g)$. Αφού η ρ είναι ανάγωγη έχουμε ότι $\deg \rho = 1$. $\square$

Πόρισμα IX.1.15. Έστω $\rho : G \rightarrow GL_k(\mathbb{F})$ ανάγωγη αναπαράσταση και $A = (a_{ij}) \in \mathbb{F}^{k,k}$. Θέτουμε

$$A_0 = \sum_{t \in G} \rho(t^{-1}) A \rho(t).$$

Τότε $A_0 = \lambda \mathbb{I}_k$, $\lambda \in \mathbb{F}$ και $k\lambda = |G| \text{tr} A$.

Απόδειξη. Παρατηρούμε ότι για κάθε $g \in G$ ισχύει ότι

$$\rho(g) A_0 = A_0 \rho(g)$$

και επομένως $A_0 = \lambda \mathbb{I}_k$. Επίσης,

$$k\lambda = \text{tr}(\lambda \mathbb{I}_k) = \text{tr}(A_0) = \sum_{t \in G} \text{tr}(\rho(t^{-1}) A \rho(t)) = |G| \text{tr}(A).$$

$\square$

Πόρισμα IX.1.16. Αν θέσουμε $(a_{ij}(g)) = \rho(g)$ τότε

$$k \sum_{t \in G} a_{ij}(t) a_{l,k}(t^{-1}) = g \delta_{i,k} \delta_{jl}.$$

Απόδειξη. Θεωρούμε τον πίνακα $A = (a_{\nu,\mu})$ με $a_{\nu,\mu} = \delta_{j,\nu} \delta_{\mu,l}$, και υπολογίζουμε τον πίνακα της

$$(c_{ij}) = \sum_{t \in G} \rho(t^{-1}) A \rho(t)$$

όπου

$$\begin{aligned} c_{i,k} &= \sum_{t \in G} \sum_{\nu=1}^k \sum_{\mu=1}^k a_{i,\nu}(t^{-1}) a_{\nu,\mu} a_{\mu,k}(t) = \sum_{t \in G} \sum_{\nu=1}^k \sum_{\mu=1}^k a_{i,\nu}(t^{-1}) \delta_{j,\nu} \delta_{\mu,l} a_{\mu,k}(t) \\ &= \sum_{t \in G} a_{i,j}(t^{-1}) a_{l,k}(t) \end{aligned}$$

Παρατηρούμε ότι $c_{i,k} = \lambda \delta_{i,k}$, ενώ η τιμή του $\lambda = |G|/k \text{tr} A = |G|/k \delta_{j,l}$, από όπου προκύπτει η ζητούμενη ιδιότητα. $\square$

Θεωρούμε το σύνολο των συναρτήσεων $G \rightarrow \mathbb{F}$ το οποίο θα το συμβολίζουμε με $G^{\mathbb{F}}$. Το σύνολο αυτό αποτελεί διανυσματικό χώρο.

Ορισμός ΙΧ.1.17. Για $\phi, \psi : G \rightarrow \mathbb{F}$ ορίζουμε

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{t \in G} \phi(t) \psi(t^{-1}).$$

Παρατήρηση ΙΧ.1.18. Η συνάρτηση $\langle \cdot, \cdot \rangle : G^{\mathbb{F}} \times G^{\mathbb{F}} \rightarrow \mathbb{F}$ είναι $\mathbb{F}$ -διγραμμική, δηλαδή:

$$\begin{aligned} \langle \lambda_1 \phi_1 + \lambda_2 \phi_2, \psi \rangle &= \lambda_1 \langle \phi_1, \psi \rangle + \lambda_2 \langle \phi_2, \psi \rangle, \quad \lambda_1, \lambda_2 \in \mathbb{F} \\ \langle \phi, \lambda_1 \psi_1 + \lambda_2 \psi_2 \rangle &= \lambda_1 \langle \phi, \psi_1 \rangle + \lambda_2 \langle \phi, \psi_2 \rangle, \quad \lambda_1, \lambda_2 \in \mathbb{F} \end{aligned}$$

Παρατήρηση ΙΧ.1.19. Στην περίπτωση $\mathbb{F} = \mathbb{C}$ μπορούμε να ορίσουμε

$$\langle \phi | \psi \rangle = \frac{1}{|G|} \sum_{t \in G} \phi(t) \overline{\psi(t)}.$$

Το τελευταίο είναι ένα ερμητιανό εσωτερικό γινόμενο. Επίσης για ένα χαρακτήρα μιας πεπερασμένης ομάδας G έχουμε ότι αν n είναι η τάξη ενός στοιχείου $g \in G$, τότε $\rho(g)^n = \rho(g^n) = \mathbb{I}_k$ συνεπώς το ελάχιστο πολυώνυμο του $\rho(g)$ διαιρεί το $x^n - 1$, έχει απλές ρίζες, άρα ο πίνακας $\rho(g)$ είναι διαγωνοποιήσιμος και έχει ως ιδιοτιμές n -ρίζες της μονάδας. Άρα υπάρχει αντιστρέψιμος πίνακας Q με

$$Q^{-1} \rho(g) Q = \text{diag}(\lambda_1, \dots, \lambda_k) \text{ και } Q^{-1} \rho(g^{-1}) Q = \text{diag}(\lambda_1^{-1}, \dots, \lambda_k^{-1}).$$

Έχουμε ότι

$$\text{tr}(\rho(g)) = \lambda_1 + \dots + \lambda_k$$

και

$$\text{tr}(\rho(g^{-1})) = \lambda_1^{-1} + \dots + \lambda_k^{-1} = \bar{\lambda}_1 + \dots + \bar{\lambda}_k.$$

Δηλαδή $\chi(g^{-1}) = \overline{\chi(g)}$. Αυτό σημαίνει ότι για χαρακτήρες χ_1, χ_2 έχουμε $\langle \chi_1, \chi_2 \rangle = \langle \chi_1 | \chi_2 \rangle$.

Θεώρημα ΙΧ.1.20. Αν χ χαρακτήρας ανάγωγης παράστασης τότε

$$\langle \chi, \chi \rangle = 1$$

Αν ϕ, ψ χαρακτήρες μη-ισοδύναμων ανάγωγων παραστάσεων τότε

$$\langle \phi, \psi \rangle = 0.$$

Απόδειξη. Θέτουμε, $(a_{ij}(t)) = \rho(t)$. Από το πόρισμα **ΙΧ.1.16** έχουμε ότι

$$\sum_{t \in G} a_{ij}(t) a_{l, \kappa}(t^{-1}) = \delta_{i, \kappa} \delta_{jl} \frac{|G|}{k}.$$

Συνεπώς,

$$\langle \chi, \chi \rangle = \frac{1}{|G|} \sum_{t \in G} \chi(t) \chi(t^{-1}) = \frac{1}{|G|} \sum_{i=1}^k a_{ii}(t) \sum_{j=1}^k a_{jj}(t^{-1}) = \frac{1}{k} \sum_{i,j=1}^k \delta_{ij} \delta_{ij} = 1.$$

Έστω τώρα $\phi = \chi_{\rho_1}$ και $\psi = \chi_{\rho_2}$ και $\rho_1(t) = (a_{ij}(t))$, $\rho_2(t) = (b_{ij}(t))$. Από την εξίσωση **ΙΧ.1** έχουμε

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{t \in G} \phi(t) \psi(t^{-1}) = \frac{1}{|G|} \sum_{t \in G} \sum_{i=1}^{k_2} a_{ii}(t) \sum_{j=1}^{k_1} b_{jj}(t^{-1}) = \frac{1}{|G|} \sum_{i,j}^{k_1, k_2} (a_{ii}(t) b_{jj}(t^{-1})) = 0.$$

□

Θεώρημα IX.1.21. Έστω $\rho : G \rightarrow GL(V)$ μια αναπαράσταση της G με χαρακτήρα ϕ και υποθέτουμε ότι διασπάται σε άθροισμα αναγωγών αναπαραστάσεων

$$V = W_1 \oplus \cdots \oplus W_s.$$

Τότε το πλήθος των W_i που είναι ισόμορφες με την ανάγωγη αναπαράσταση W με χαρακτήρα χ είναι ίσο με το $\langle \phi, \chi \rangle$.

Απόδειξη. Αν χ_i είναι ο χαρακτήρας της W_i τότε $\phi = \chi_1 + \cdots + \chi_s$, οπότε

$$\langle \phi, \chi \rangle = \langle \phi, \chi_1 \rangle + \cdots + \langle \phi, \chi_s \rangle$$

και το ζητούμενο προκύπτει από το θεώρημα IX.1.20. □

Πόρισμα IX.1.22. Δύο αναπαραστάσεις με τον ίδιο χαρακτήρα είναι ισόμορφες.

Απόδειξη. Πράγματι, από το θεώρημα IX.1.21 αν δύο αναπαραστάσεις έχουν τους ίδιους χαρακτήρες έχουν την ίδια ανάλυση ως άθροισμα αναγωγών αναπαραστάσεων. □

Αν μια αναπαράσταση $\rho : G \rightarrow GL(V)$ έχει μια διάσπαση σε άθροισμα αναπαραστάσεων

$$V = m_1 V_1 \oplus \cdots \oplus m_s V_s$$

όπου W_i δεν είναι ισοδύναμη με την W_j για $i \neq j$, και ϕ είναι ο χαρακτήρας της ρ , τότε υπολογίζουμε ότι

$$\langle \phi, \phi \rangle = \sum_{i=1}^s m_i^2 \in \mathbb{N}.$$

Από την παραπάνω έκφραση έχουμε ότι ϕ είναι ανάγωγη αν και μόνο αν $\langle \phi, \phi \rangle = 1$.

Πρόταση IX.1.23. Κάθε ανάγωγη αναπαράσταση $\rho_i : G \rightarrow GL(W_i)$ εμφανίζεται ως ευθύς προσδετός της ομαλής αναπαράστασης με πολυπλοκότητα ίση με τον $\deg \rho_i$.

Απόδειξη. Υπολογίζουμε

$$\langle r_G, \chi_i \rangle = \frac{1}{|G|} \sum_{t \in G} r_G(t^{-1}) \chi_i(t) = \frac{|G|}{|G|} \chi_i(1) = \deg \rho_i.$$

□

Πόρισμα IX.1.24. Ισχύει ότι

- $\sum_{i=1}^s \deg^2 \rho_i = |G|$
- Αν $g \neq e_G$ τότε $\sum_{i=1}^s \deg \rho_i \chi_i(g) = 0$.

Ορισμός ΙΧ.1.25. Μια συνάρτηση $f : G \rightarrow \mathbb{F}$ θα λέγεται συνάρτηση κλάσεων συζυγίας αν για κάθε $g, h \in G$ ισχύει $f(hgh^{-1}) = f(g)$, δηλαδή αν $f(gh) = f(hg)$.

Παρατήρηση ΙΧ.1.26. Ο χώρος $\text{cf}(G)$ των συναρτήσεων κλάσεων συζυγίας είναι ένας $\mathbb{F}$ -διανυσματικός χώρος, υπόχωρος του χώρου των συναρτήσεων $f : G \rightarrow \mathbb{F}$. Οι χαρακτήρες αναπαράστασεων ανήκουν στον χώρο κλάσεων συζυγίας. Η διγραμμική μορφή $\langle \cdot, \cdot \rangle$ περιορισμένη στο $\text{cf}(G)$ δίνει ένα εσωτερικό γινόμενο στο $\text{cf}(G)$.

Πρόταση ΙΧ.1.27. Έστω f μια συνάρτηση κλάσεων στην G και $\rho : G \rightarrow \text{GL}(V)$ μια αναπαράσταση της G . Θεωρούμε την συνάρτηση

$$\rho_f = \sum_{t \in G} f(t)\rho(t).$$

Αν η ρ είναι ανάγωγη βαθμού n και χαρακτήρα χ , τότε

$$\rho_f = \lambda \text{Id}_V, \text{ με } \lambda = \frac{1}{n} \sum_{t \in G} f(t)\chi(t) = \frac{|G|}{n} \langle f | \bar{\chi} \rangle.$$

Απόδειξη. Υπολογίζουμε

$$\rho(g)^{-1} \rho_f \rho(g) = \sum_{t \in G} f(t) \rho(g)^{-1} \rho(t) \rho(g) = \sum_{t \in G} f(t) \rho(g^{-1}tg)$$

Στον παραπάνω τύπο θέτουμε $h = g^{-1}tg$ και έχουμε

$$\rho(g)^{-1} \rho_f \rho(g) = \sum_{u \in G} f(hg^{-1}g) \rho(u) = \sum_u f(u) \rho(u) = \rho_f.$$

Συνεπώς, $\rho_f \rho(g) = \rho(g) \rho_f$ από όπου έχουμε ότι $\rho_f = \lambda \text{Id}_V$. Το τελευταίο έχει ίχνος $n\lambda$ ενώ το ίχνος του ρ_f υπολογίζεται σε

$$\text{tr}(\rho_f) = \sum_{t \in G} f(t)\chi(t) = |G| \langle f | \bar{\chi} \rangle,$$

από όπου προκύπτει το ζητούμενο. □

Θεώρημα ΙΧ.1.28. Η διάσταση του $\text{cf}(G)$ είναι ίση με το πλήθος των κλάσεων συζυγίας. Οι ανάγωγοι χαρακτήρες αποτελούν μια ορθοκανονική βάση του χώρου $\text{cf}(G)$.

Απόδειξη. Η σχέση $g \sim g'$ αν και μόνο αν $g' = hgh^{-1}$ είναι μια σχέση ισοδυναμίας στα στοιχεία της G . Θεωρούμε τις κλάσεις ισοδυναμίας $C_1, \dots, C_{h_G}$ τις οποίες ονομάζουμε κλάσεις συζυγίας. Έστω $\{g_1, \dots, g_{h_G}\}$ ένα πλήρες σύστημα αντιπροσώπων των κλάσεων. Παρατηρούμε ότι γνωρίζουμε πλήρως μια συνάρτηση στο $\text{cf}(G)$ αν και μόνο αν την γνωρίζουμε στα $g_1, \dots, g_{h_G}$. Η συνάρτηση

$$\begin{aligned} \Phi : \text{cf}(G) &\longrightarrow \mathbb{F}^{h_G} \\ f &\longmapsto (f(g_1), \dots, f(g_{h_G})) \end{aligned}$$

είναι ισομορφισμός διανυσματικών χώρων.

Ισχύει για τους ανάγωγους χαρακτήρες ότι $\langle \chi_i, \chi_j \rangle = \delta_{ij}$. Συνεπώς οι ανάγωγοι χαρακτήρες αποτελούν ένα ορθοκανονικό σύστημα και είναι γραμμικά ανεξάρτητοι. Άρα το πλήθος τους m είναι πεπερασμένο και μικρότερο του h_G .

Θα αποδείξουμε ότι $m = h_G$ δηλαδή υπάρχουν ακριβώς h_G το πλήθος ανάγωγες αναπαραστάσεις της G .

Θα αποδείξουμε ότι αν $f \in \text{cf}(G)$ και $\langle f, \chi_i \rangle = 0$ τότε $f = 0$. Συνεπώς, το ορθοκανονικό σύστημα των ανάγωγων χαρακτήρων δεν μπορεί να επεκταθεί σε μια μεγαλύτερη ορθοκανονική βάση του χώρου των συναρτήσεων κλάσεων άρα αποτελεί βάση του χώρου αυτού.

Πράγματι, για ένα τέτοιο f και κάθε αναπαράσταση ρ της G , έχουμε

$$\rho_f = \sum_{t \in G} f(t) \rho(t).$$

Αφού το f είναι κάθετο στα $\bar{\chi}$, από την πρόταση IX.1.27 έχουμε ότι ρ_f είναι 0 αν η ρ είναι ανάγωγη. Αλλά ακόμα και αν δεν είναι ανάγωγη, την διασπάμε σε ανάγωγες αναπαραστάσεις και πάλι καταλήγουμε στο ότι $\rho_f = 0$.

Τέλος, αν πάρουμε ρ την ομαλή παράσταση και υπολογίσουμε

$$0 = \rho_f e_1 = \sum_{t \in G} f(t) \rho(t) e_1 = \sum_{t \in G} f(t) e_t$$

και αφού τα $\{e_t\}_{t \in G}$ είναι βάση, έχουμε ότι $f(t) = 0$ για κάθε $t \in G$, δηλαδή το ζητούμενο. $\square$

Παρατήρηση IX.1.29. Υπάρχουν διαφορετικές γενικεύσεις της θεωρίας αναπαραστάσεων. Ένα σημαντικό εργαλείο της θεωρίας είναι ο τελεστής μέσου όρου μιας ποσότητας πάνω στην ομάδα δηλαδή η έκφραση

$$\frac{1}{|G|} \sum_{g \in G} f(g)$$

για μια συνάρτηση $f : G \rightarrow \mathbb{F}$. Για άπειρες ομάδες που έχουν πεπερασμένο «όγκο» $\int_G 1$ μπορούμε να έχουμε μια παρόμοια θεωρία όπου τα αθροίσματα αντικαθίστανται από ολοκληρώματα δηλαδή εκφράσεις της μορφής:

$$\frac{1}{\int_G 1} \int_G f(t).$$

Παρατηρήστε ότι στην περίπτωση των πεπερασμένων ομάδων

$$|G| = \sum_{t \in G} 1 \text{ το οποίο είναι το ανάλογο της ποσότητας } \int_G 1.$$

Η αντικατάσταση αθροισμάτων με ολοκληρώματα θα σας γίνει σαφέστερη σε κάποιο μάθημα θεωρίας μέτρου ή πιθανοτήτων. Αρχικά στοιχεία μιας τέτοιας θεωρίας έχουμε ήδη δει στην θεωρία της ανάλυσης Fourier, στο παράδειγμα VII.6.7, η οποία είναι μια θεωρία αναπαραστάσεων του κύκλου S^1 ο οποίος έχει πεπερασμένο όγκο 2π .

Η θεωρία αναπαραστάσεων στην περίπτωση που είτε ο «όγκος» της ομάδας δεν είναι πεπερασμένος είτε δεν μπορούμε να αντιστρέψουμε την τάξη της ομάδας, όπως συμβαίνει σε ένα σώμα $\mathbb{F}$ χαρακτηριστικής p με $p \mid |G|$ είναι πολύ δυσκολότερη. Αυτές οι αναπαραστάσεις είναι ενδιαφέρουσες. Για παράδειγμα η συνάρτηση, η οποία σχετίζεται με τις κατασκευές στα δυναμικά συστήματα της παραγράφου VII.5.5:

$$\begin{aligned} (\mathbb{R}, +) &\longrightarrow \text{GL}_n(\mathbb{R}) \\ t &\longmapsto \exp(tA). \end{aligned} \tag{IX.2}$$

IX.2 Ένα μοντέλο της Κβαντομηχανικής

Στην παράγραφο αυτή θα δώσουμε ένα μοντέλο της κβαντομηχανικής βασισμένο στις αυτοσυζυγείς γραμμικές συναρτήσεις από ένα πεπερασμένης διάστασης διανυσματικό χώρο στον

εαυτό του. Ακολουθούμε το εξαιρετικό βιβλίο των L.D. Faddeev και O. A. Yakubovskii [6]. Ο αναγνώστης θα ωφεληθεί αν επιστρέψει και ξαναδιαβάσει την παράγραφο αυτή αφού έχει μελετήσει ένα μάθημα θεωρίας πιθανοτήτων.

Στο μοντέλο μας η άλγεβρα των παρατηρησίμων μεγεθών $\mathfrak{A}$ θα είναι η άλγεβρα των αυτοσυζυγών γραμμικών συναρτήσεων $V \rightarrow V$, όπου ο V είναι ένας πεπερασμένης διάστασης διανυσματικός χώρος πάνω από τους μιγαδικούς αριθμούς. Ο χώρος V θεωρείται εφοδιασμένος με το κανονικό ερμητιανό εσωτερικό γινόμενο, δηλαδή αν $B = \{e_1, \dots, e_n\}$ είναι μια βάση του V και $x, y \in V$, έχουν αναπαράσταση σε συντεταγμένες

$$[x]_B = (x_1, \dots, x_n)^t, \quad [y]_B = (y_1, \dots, y_n)^t,$$

το εσωτερικό γινόμενο των x, y δίνεται από το τύπο

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i.$$

Η παραπάνω επιλογή εσωτερικού γινομένου κάνει την βάση B ορθοκανονική και επιπλέον κάθε γραμμική συνάρτηση $A : V \rightarrow V$ περιγράφεται από τον πίνακα

$$A = (a_{ij}), \quad a_{ij} = \langle A e_i, e_j \rangle.$$

Ο πίνακας A είναι αυτοσυζυγής αν και μόνο αν $A = A^*$, δηλαδή αν $a_{ij} = \bar{a}_{ji}$.

Παρατηρούμε ότι αν $A, B \in \mathfrak{A}$, τότε $A + B \in \mathfrak{A}$, όπως και $\lambda A \in \mathfrak{A}$ για κάθε $\lambda \in \mathbb{R}$.

Θέλουμε να ορίσουμε συναρτήσεις στην $\mathfrak{A}$ ως εξής:

- Αν η $f(z) = \sum_{i=0}^{\infty} a_i z^i$ είναι μια αναλυτική συνάρτηση, ορίζουμε (λήμμα VII.5.4)

$$f(A) = \sum_{i=0}^{\infty} a_i A^i$$

- Αφού η A είναι αυτοσυζυγής είναι διαγωνοποιήσιμη και έχουμε μια βάση από ιδιοδιανύσματα $\{\phi_1, \dots, \phi_n\}$, ώστε

$$A \phi_i = \lambda_i \phi_i.$$

Από το φασματικό θεώρημα VIII.3.5 αρκεί να ορίσουμε την f πάνω στις ιδιοδιανύσματα, δηλαδή

$$f(A) \phi_i = f(\lambda_i) \phi_i,$$

δηλαδή ως μια διαγωνοποιήσιμη συνάρτηση με ιδιοτιμές $f(\lambda_i)$ στα ιδιοδιανύσματα ϕ_i .

Θα ορίσουμε την «σύνθεση» $A \circ B$ ως εξής:

$$A \circ B = \frac{(A+B)^2 - (A-B)^2}{4} = \frac{AB+BA}{2}.$$

Είναι σαφές ότι αν $A, B \in \mathfrak{A}$ τότε και $A \circ B \in \mathfrak{A}$.

Τέλος ορίζουμε την αγκύλη Lie ως εξής:

$$[A, B] = AB - BA$$

η οποία έχει τις παρακάτω ιδιότητες:

1. $[A, B] = -[B, A]$,
2. $[A + \lambda B, C] = [A, C] + \lambda [A, B]$,
3. $[A, B \circ C] = [A, B] \circ C + B \circ [A, C]$,

$$4. [A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0,$$

$$5. [A, BC] = [A, B]C + B[A, C]$$

Παρατηρούμε ότι η $[A, B]$ δεν είναι αυτοσυζυγής, αντίθετα η παράσταση $\frac{i}{\hbar}[A, B]$ είναι αυτοσυζυγής. Πράγματι,

$$\begin{aligned} [A, B]^* &= (AB - BA)^* = BA - AB \\ \left(\frac{i}{\hbar}[A, B]\right)^* &= -\frac{i}{\hbar}(AB - BA)^* = \frac{i}{\hbar}(AB - BA) = \frac{i}{\hbar}[A, B]. \end{aligned}$$

Η σταθερά $\hbar$ προσδιορίζεται από τα πειραματικά δεδομένα και προσδιορίζεται ίση με την σταθερά του Planck $\hbar = 6.62607015 \times 10^{-34} \text{ J Hz}^{-1} = 4.135667696 \times 10^{-15} \text{ eV Hz}^{-1}$.

Ορίζουμε

$$\{A, B\} = \frac{i}{\hbar}[A, B].$$

Ορισμός IX.2.1. Μία κατάσταση (state) ω στην άλγεβρα $\mathfrak{A}$ είναι ένα στοιχείο του δυικού χώρου $\mathfrak{A}^*$ δηλαδή μια γραμμική συνάρτηση

$$\begin{aligned} \omega : \mathfrak{A} &\longrightarrow \mathbb{R}, \\ A &\longmapsto \langle A | \omega \rangle \end{aligned}$$

το οποίο επιπλέον ικανοποιεί τις ιδιότητες

1. $\langle \lambda A + B | \omega \rangle = \lambda \langle A | \omega \rangle + \langle B | \omega \rangle,$
2. $\langle A^2 | \omega \rangle \geq 0,$
3. $\langle c \mathbb{I}_n | \omega \rangle = c$ για κάθε $c \in \mathbb{R},$
4. $\overline{\langle A, \omega \rangle} = \langle A, \omega \rangle.$

Η γενική μορφή ενός συναρτησοειδούς είναι

$$\langle A | \omega \rangle = \text{tr} M A.$$

Παρατηρούμε ότι ο $n \times n$ πίνακας M οδηγεί σε μια κατάσταση ω αν ικανοποιεί τις σχέσεις:

1. $M^* = M,$
2. $\langle Mx, x \rangle \geq 0$ για κάθε $x \in \mathbb{C}^n,$
3. $\text{tr} M = 1.$

Πράγματι αφού $\text{tr} M A \in \mathbb{R}$, έχουμε ότι

$$\begin{aligned} \overline{\text{tr} M A} &= \sum_{i,k=1}^n \overline{A_{ik} M_{ki}} = \sum_{i,k=1}^n A_{ki}^* M_{ik}^* \\ &= \sum_{i,k=1}^n A_{ki} M_{ik}^* = \text{tr} A M^* = \text{tr} A M. \end{aligned}$$

και αφού η παραπάνω ιδιότητα ισχύει για κάθε A έχουμε $M^* = M.$

Χρησιμοποιούμε τώρα την ιδιότητα

$$\langle A^2 | \omega \rangle = \text{tr} A^2 M \geq 0$$

Θεωρούμε την γραμμική συνάρτηση προβολής P_η στο διάνυσμα η , με $\|\eta\| = 1$ δηλαδή

$$P_\eta \xi = \langle \xi, \eta \rangle \eta.$$

Για $A = P_\eta$, θεωρούμε την ορθοκανονική βάση $\{e_1 = \eta, e_2, \dots, e_n\}$ και υπολογίζουμε ως προς αυτή την βάση

$$\text{tr}(P_\eta^2 M) = \text{tr}(P_\eta M) = \sum_{i=1}^n \langle P_\eta M e_i, e_i \rangle = \langle M \eta, \eta \rangle \geq 0.$$

δηλαδή το να είναι πίνακας M θετικά ορισμένος είναι αναγκαία συνθήκη για να ισχύει $\langle A^2 | \omega \rangle \geq 0$. Για να δείξουμε ότι είναι και ικανή παρατηρούμε ότι

$$\text{tr}(A^2 M) = \text{tr}(A M A) = \sum_{i=1}^n \langle A M A e_i, e_i \rangle = \sum_{i=1}^n \langle M A e_i, A e_i \rangle \geq 0.$$

Τέλος η συνθήκη $\text{tr}(M) = 1$ προκύπτει από την συνθήκη 3. για $c = 1$.

Ορισμός ΙΧ.2.2. Ο $n \times n$ πίνακας M που ικανοποιεί τις παραπάνω ιδιότητες 1-3 ονομάζεται πίνακας πυκνότητας της κατάστασης ω .

Πρόταση ΙΧ.2.3. Υπάρχει μια ένα προς ένα και επί αντιστοιχία ανάμεσα στις καταστάσεις ω και στο σύνολο των πινάκων πυκνότητας.

Ορισμός ΙΧ.2.4. Για δύο καταστάσεις ω_1, ω_2 ορίζουμε τον κυρτό συνδυασμό τους

$$\omega = a\omega_1 + (1-a)\omega_2 \text{ για κάποιο } 0 < a < 1. \quad (\text{IX.3})$$

Καταστάσεις οι οποίες δεν εκφράζονται ως κυρτοί συνδυασμοί άλλων καταστάσεων ονομάζονται καθαρές (pure) αλλιώς ονομάζονται μικτές (mixed).

Παρατηρούμε ότι αν η κατάσταση ω_1 αντιστοιχεί στον πίνακα M_1 και η κατάσταση ω_2 στον πίνακα M_2 στην μικτή κατάσταση ω της (IX.3) αντιστοιχεί ο πίνακας

$$M = aM_1 + (1-a)M_2.$$

Λήμμα ΙΧ.2.5. Οι πίνακες των γραμμικών συναρτήσεων προβολών P_ψ για $\|\psi\| = 1$ αποτελούν πίνακες πυκνότητας καταστάσεων.

Απόδειξη. Παρατηρούμε ότι ισχύουν οι ιδιότητες:

1. $\langle P_\psi \xi, \eta \rangle = \langle \xi, \psi \rangle \langle \psi, \eta \rangle = \overline{\langle \eta, \psi \rangle} \langle \xi, \psi \rangle = \langle \xi, P_\psi \eta \rangle.$
2. $\langle P_\psi \xi, \xi \rangle = \langle \xi, \psi \rangle \langle \psi, \xi \rangle = |\langle \psi, \xi \rangle|^2 \geq 0.$
3. $\text{tr}(P_\psi) = \langle \psi, \psi \rangle = 1.$

□

Ορισμός IX.2.6. Το διάνυσμα $\xi \in \mathbb{C}^n$, $\|\xi\| = 1$ λέγεται ένα διάνυσμα κατάστασης.

Λήμμα IX.2.7. Θεωρούμε ένα ερμητιανό θετικό πίνακα A . Για κάθε δύο διανύσματα ϕ, ψ ισχύει

$$|\langle A\phi, \psi \rangle|^2 \leq \langle A\phi, \phi \rangle \langle A\psi, \psi \rangle. \quad (\text{IX.4})$$

Απόδειξη. Πράγματι, έχουμε ότι για κάθε $a, b \in \mathbb{C}$, έχουμε

$$\begin{aligned} 0 &\leq \langle A(a\phi + b\psi), a\phi + b\psi \rangle = a\bar{a}\langle A\phi, \phi \rangle + a\bar{b}\langle A\phi, \psi \rangle + b\bar{a}\langle A\psi, \phi \rangle + b\bar{b}\langle A\psi, \psi \rangle \\ &= \chi^* \begin{pmatrix} \langle A\phi, \phi \rangle & \langle A\phi, \psi \rangle \\ \langle A\psi, \phi \rangle & \langle A\psi, \psi \rangle \end{pmatrix} \chi \end{aligned}$$

όπου $\chi = (a, b)^t$. Δηλαδή ο πίνακας

$$\begin{pmatrix} \langle A\phi, \phi \rangle & \langle A\phi, \psi \rangle \\ \langle A\psi, \phi \rangle & \langle A\psi, \psi \rangle \end{pmatrix}$$

είναι θετικά ορισμένος και έχει θετική ορίζουσα $\langle A\phi, \phi \rangle \langle A\psi, \psi \rangle - \langle A\phi, \psi \rangle \overline{\langle A\phi, \psi \rangle} \geq 0$ από όπου προκύπτει η ζητούμενη ανισότητα. $\square$

Πρόταση IX.2.8. Όλες οι καταστάσεις οι οποίες αντιστοιχούν στους πίνακες της γραμμικής συνάρτησης P_ψ είναι καθαρές.

Απόδειξη. Θα πρέπει να δείξουμε ότι η ισότητα

$$P_\psi = aM_1 + (1-a)M_2, 0 < a < 1, \quad (\text{IX.5})$$

δίνει ότι $P_\psi = M_1 = M_2$.

Από την (IX.4) έχουμε ότι αν $\langle A\phi, \phi \rangle = 0$ τότε $A\phi = 0$.

Θεωρούμε τον χώρο H_1 των διανυσμάτων που είναι κάθετα στο ψ , δηλαδή $P_\psi \phi = 0$ για κάθε $\phi \in H_1$. Από την εξίσωση (IX.5) έχουμε

$$0 \leq a\langle M_1\phi, \phi \rangle \leq a\langle M_1\phi, \phi \rangle + (1-a)\langle M_2\phi, \phi \rangle = \langle P_\psi\phi, \phi \rangle = 0,$$

συνεπώς $M_1\phi = 0$ για κάθε $\phi \in H_1$. Αφού η M_1 είναι αυτοσυζυγής έχουμε για κάθε $\xi \in \mathbb{C}^n$

$$\langle M_1\phi, \xi \rangle = \langle \phi, M_1\xi \rangle = 0,$$

συνεπώς $M_1\xi = c_\xi\psi$, όπου $c_\xi \in \mathbb{C}$. Ιδιαίτερα ισχύει $M_1\psi = c_\psi\psi$. Το τυχαίο διάνυσμα ξ γράφεται ως

$$\xi = \langle \xi, \psi \rangle \psi + \phi, \text{ όπου } \phi \in H_1.$$

Συνεπώς

$$M_1\xi = c_\psi P_\psi \xi,$$

δηλαδή $M_1 = c_\psi P_\psi$ και η συνθήκη $\text{tr}(M_1) = 1$ επιβάλλει ότι $c_\psi = 1$. Άρα $M_1 = M_2 = P_\psi$. $\square$

Πρόταση ΙΧ.2.9. Κάθε κατάσταση είναι ένας κυρτός συνδυασμός από καθαρές καταστάσεις.

Απόδειξη. Θεωρούμε ένα πίνακα πυκνότητας M ο οποίος είναι εξορισμού αυτοσυζυγής, συνεπώς έχει μια ορθοκανονική βάση $\{\phi_1, \dots, \phi_n\}$ αποτελούμενη από ιδιοδιανύσματα,

$$M\phi_i = \mu_i\phi_i.$$

Το τυχαίο διάνυσμα γράφεται ως γραμμικός συνδυασμός στοιχείων της βάσης αυτής:

$$\xi = \sum_{i=1}^n \langle \xi, \phi_i \rangle \phi_i.$$

Υπολογίζουμε ότι

$$M\xi = \sum_{i=1}^n \mu_i \langle \xi, \phi_i \rangle \phi_i = \sum_{i=1}^n \mu_i P_{\phi_i} \xi,$$

συνεπώς

$$M = \sum_{i=1}^n \mu_i P_{\phi_i}.$$

Όλες οι ιδιοτιμές είναι μ_i είναι μη-αρνητικές αφού είναι ιδιοτιμές ενός θετικού πίνακα. Επίσης

$$\sum_{i=1}^n \mu_i = \text{tr}(M) = 1$$

και συνεπώς το M είναι κυρτός συνδυασμός των καθαρών καταστάσεων P_{ϕ_i} . □

Πόρισμα ΙΧ.2.10. Αν η κατάσταση ω που αντιστοιχεί στον πίνακα M είναι καθαρή, τότε $M = P_\psi$ για κάποιο διάνυσμα ψ , μέτρου 1.

Απόδειξη. Από την απόδειξη της πρότασης ΙΧ.2.9 έχουμε ότι

$$M = \sum_{i=1}^n \mu_i P_{\phi_i},$$

συνεπώς έχουμε ότι μόνο μία ιδιοτιμή $\mu_1 = 1$ και όλες οι άλλες ιδιοτιμές είναι μηδενικές, δηλαδή $M = P_{\phi_1}$. □

ΙΧ.2.1 Οι σχέσεις αβεβαιότητας του Heisenberg

Ο υπολογισμός $\langle A|\omega \rangle \in \mathbb{R}$ αποκτά την φυσική σημασία της μέσης τιμής του παρατηρήσιμου A στην κατάσταση ω οπότε θέτουμε

$$A_{\text{avg}} = \langle A, \omega \rangle.$$

Ορίζουμε

$$\Delta_\omega^2 A = \langle (A - A_{\text{avg}})^2 | \omega \rangle = \langle A^2 | \omega \rangle - \langle A | \omega \rangle^2$$

και θέτουμε

$$\Delta_\omega A = \sqrt{\Delta_\omega^2 A}.$$

Οι σχέσεις αβεβαιότητας εκφράζονται ως

Πρόταση IX.2.11.

$$\Delta_\omega A \cdot \Delta_\omega B \geq \frac{\hbar}{2} |\langle \{A, B\}_h, \omega \rangle|. \quad (\text{IX.6})$$

Απόδειξη. Θα αποδείξουμε το ζητούμενο πρώτα για καθαρές καταστάσεις.

Ισχύει ότι

$$\langle (A + iaB)\psi, (A + iaB)\psi \rangle \geq 0, \text{ για κάθε } \psi \in \mathbb{C}^n, \|\psi\| = 1, a \in \mathbb{R}.$$

Επεκτείνουμε αριστερό μέρος

$$\langle A^2\psi, \psi \rangle + a^2 \langle B^2\psi, \psi \rangle + ia \langle (AB - BA)\psi, \psi \rangle \geq 0.$$

το οποίο το γράφουμε ως

$$\langle A^2|\omega \rangle + a^2 \langle B^2|\omega \rangle - ah \langle \{AB - BA\}_h|\omega \rangle \geq 0$$

η οποία ισχύει για κάθε $a \in \mathbb{R}$ από όπου προκύπτει ότι

$$\langle A^2|\omega \rangle \langle B^2|\omega \rangle \geq \frac{\hbar^2}{4} \langle \{AB - BA\}_h|\omega \rangle^2.$$

Το ζητούμενο προκύπτει με την αντικατάσταση $A \rightarrow A - A_{\text{avg}}, B \rightarrow B - B_{\text{avg}}$ και την σχέση $\{A, B\}_h = \{A - A_{\text{avg}}, B - B_{\text{avg}}\}_h$.

Για μικτές καταστάσεις έχουμε ότι η διασπορά μεγαλώνει και οι ανισότητες ανάγονται στις αντίστοιχες ανισότητες για καθαρές καταστάσεις. Πράγματι, αν

$$\omega = a\omega_1 + (1 - a)\omega_2$$

έχουμε ότι

$$\begin{aligned} \Delta_\omega^2(A) &= \langle A^2|\omega \rangle - \langle A|\omega \rangle^2 \\ &= a \langle A^2|\omega_1 \rangle + (1 - a) \langle A^2|\omega_2 \rangle - (a \langle A|\omega_1 \rangle + (1 - a) \langle A|\omega_2 \rangle)^2 \\ &= a \Delta_{\omega_1}^2(A) + (1 - a) \Delta_{\omega_2}^2(A) + a(a - 1) (\langle A|\omega_1 \rangle^2 + \langle A|\omega_2 \rangle^2) - 2a(1 - a) \langle A|\omega_1 \rangle \langle A|\omega_2 \rangle \\ &= a \Delta_{\omega_1}^2(A) + (1 - a) \Delta_{\omega_2}^2(A) - a(a - 1) (\langle A|\omega_1 \rangle - \langle A|\omega_2 \rangle)^2 \\ &\geq a \Delta_{\omega_1}^2(A) + (1 - a) \Delta_{\omega_2}^2(A) \end{aligned}$$

με την ισότητα να ισχύει μόνο αν $\langle A|\omega_1 \rangle = \langle A|\omega_2 \rangle$. Έχουμε τώρα

$$\begin{aligned} \Delta_\omega^2 A \cdot \Delta_\omega^2 B &\geq (a \Delta_{\omega_1}^2(A) + (1 - a) \Delta_{\omega_2}^2(A)) (a \Delta_{\omega_1}^2(B) + (1 - a) \Delta_{\omega_2}^2(B)) \\ &\geq (a \Delta_{\omega_1} A \cdot \Delta_{\omega_1} B + (1 - a) \Delta_{\omega_2} A \cdot \Delta_{\omega_2} B)^2 \end{aligned}$$

από όπου καταλήγουμε στην

$$\Delta_\omega A \cdot \Delta_\omega B \geq a \Delta_{\omega_1} A \cdot \Delta_{\omega_1} B + (1 - a) \Delta_{\omega_2} A \cdot \Delta_{\omega_2} B$$

με ισότητα να ισχύει αν και μόνο αν $\langle A|\omega_1 \rangle = \langle A|\omega_2 \rangle$ και $\langle B|\omega_1 \rangle = \langle B|\omega_2 \rangle$. □

Οι σχέσεις αβεβαιότητας δείχνουν ότι η διασπορές δύο παρατηρήσιμων, μπορεί να είναι και οι δύο μηδέν αν η μέση τιμή του $\{A, B\}_h$ είναι 0. Για παρατηρήσιμα A, B που αντιμετωπίζονται έχουμε ότι $\{A, B\}_h = 0$ και η μέση τιμή είναι μηδέν για όλες τις καταστάσεις ω . Σε αυτή την περίπτωση οι σχέσεις αβεβαιότητας δεν βάζουν κανένα περιορισμό και τα παρατηρήσιμα A, B είναι ταυτόχρονα μετρήσιμα.

Αντιθέτως αν $\{A, B\}_h = I_n$, όπως συμβαίνει με τα παρατηρήσιμα μεγέθη θέσης, ορμής, τότε οι σχέσεις αβεβαιότητας παίρνουν την ισχυρότερη μορφή τους

$$\Delta_\omega A \cdot \Delta_\omega B \geq \frac{\hbar}{2},$$

δηλαδή δεν υπάρχουν καταστάσεις ω στις οποίες τα A, B να είναι ταυτόχρονα μετρήσιμα.

Η φυσική ερμηνεία των ιδιοτιμών και ιδιοδιανυσμάτων των παρατηρήσιμων

Θεωρούμε την συνάρτηση Heaviside

$$\theta(x) = \begin{cases} 1 & \text{αν } x \geq 0, \\ 0 & \text{αν } x < 0. \end{cases}$$

Η παραπάνω συνάρτηση δεν είναι αναλυτική και συνεπώς οποιοσδήποτε υπολογισμός αυτής γίνει πάνω σε ένα παρατηρήσιμο μέγεθος θα πρέπει να γίνει με την την βοήθεια του φασματικού θεωρήματος.

Η έκφραση

$$\omega_A(\lambda) = \langle \theta(\lambda \mathbb{I}_n - A) | \omega \rangle$$

είναι μια συνάρτηση κατανομής και εκφράζει την πιθανότητα μια μέτρηση $\langle A | \omega \rangle$ του παρατηρήσιμου $A \in \mathfrak{A}$ στην κατάσταση ω να είναι $\leq \lambda$. Για να υπολογίσουμε την τιμή $\theta(\lambda \mathbb{I}_n - A) \in \mathfrak{A}$ θεωρούμε τα ιδιοδιανύσματα και τις ιδιοτιμές του A ,

$$A\phi_i = \lambda_i \phi_i.$$

Αρχικά θα θεωρήσουμε την περίπτωση που οι ιδιοτιμές $\lambda_1, \dots, \lambda_n$ είναι διαφορετικές και θα τις διατάξουμε σε αύξουσα σειρά

$$\lambda_1 < \lambda_2 < \dots < \lambda_n.$$

Ας θυμηθούμε ότι οι ιδιοτιμές μιας αυτοσυζυγούς γραμμικής συνάρτησης είναι όλες πραγματικές. Θεωρούμε τις συναρτήσεις προβολής P_{ϕ_i} στους διαφορετικούς ιδιοχώρους. Θεωρούμε την συνάρτηση

$$P_A(\lambda) = \sum_{\lambda_i \leq \lambda} P_{\phi_i}.$$

Θα δείξουμε ότι

$$P_A(\lambda) = \theta(\lambda \mathbb{I}_n - A).$$

Για αυτό αρκεί να δείξουμε ότι οι $P_A(\lambda)$ και $\theta(\lambda \mathbb{I}_n - A)$ δρουν με τον ίδιο τρόπο πάνω στα ιδιοδιανύσματα $\phi_1, \dots, \phi_n$. Έχουμε ότι

$$\theta(\lambda \mathbb{I}_n - A)\phi_j = \theta(\lambda - \lambda_j)\phi_j = \begin{cases} \phi_j & \text{αν } \lambda \geq \lambda_j \\ 0 & \text{αν } \lambda < \lambda_j \end{cases}$$

Από την άλλη

$$P_A(\lambda)\phi_j = \sum_{\lambda_i \leq \lambda} P_{\phi_i}\phi_j = \begin{cases} \phi_j & \text{αν } \lambda \geq \lambda_j \\ 0 & \text{αν } \lambda < \lambda_j \end{cases}$$

δηλαδή το ζητούμενο.

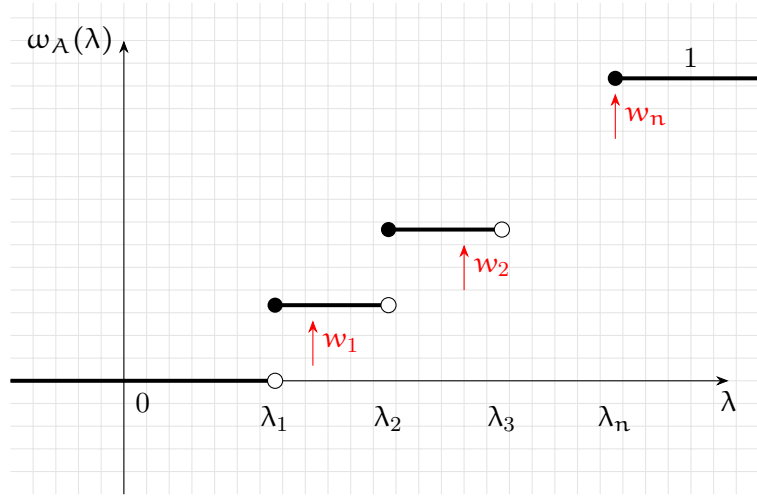
Η συνάρτηση $P_A(\lambda)$ είναι προβολή, δηλαδή $P_A(\lambda)^* = P_A(\lambda)$ και $P_A(\lambda)^2 = P_A(\lambda)$. Είναι σαφές ότι $P_A(\lambda) = 0$ αν $\lambda < \lambda_1$ και $P_A(\lambda) = 1$ αν $\lambda \geq \lambda_n$. Η συνάρτηση $P_A(\lambda)$ ονομάζεται η φασματική συνάρτηση του A . Αν η κατάσταση ω αντιστοιχεί στον πίνακα πυκνότητας M τότε

$$\omega_A(\lambda) = \langle P_A(\lambda) | \omega \rangle = \text{tr}(MP_A(\lambda)) = \text{tr}\left(M \sum_{\lambda_i \leq \lambda} P_{\phi_i}\right) = \sum_{\lambda_i \leq \lambda} \text{tr}(MP_{\phi_i}),$$

οπότε καταλήγουμε στην

$$\omega_A(\lambda) = \sum_{\lambda_i \leq \lambda} \langle M\phi_i, \phi_i \rangle.$$

Υπενθυμίζουμε ότι η ποσότητες $\langle M\phi_i, \phi_i \rangle \geq 0$ δηλαδή η $\omega_A(\lambda)$ είναι μια κατά τμήματα σταθερή συνάρτηση με πηδήματα στις τιμές λ που είναι οι ιδιοτιμές λ_i και στα σημεία αυτά είναι δεξιά συνεχής. Το γράφημα της παρουσιάζεται στο σχήμα IX.1.

Σχήμα IX.1: Η συνάρτηση κατανομής $\omega_A(\lambda)$

Η πιθανότητα w_i να πάρουμε την τιμή λ_i είναι ίση με το πήδημα της $\omega_A(\lambda)$ στο σημείο αυτό δηλαδή

$$w_i = \langle M\phi_i, \phi_i \rangle.$$

Το άθροισμα όλων των w_i είναι ίσο με 1 αφού

$$\sum_{i=1}^n w_i = \text{tr}(M) = 1.$$

Για μία καθαρή κατάσταση με πίνακα P_ψ , έχουμε ότι

$$w_i = |\langle \psi, \phi_i \rangle|^2.$$

Τέλος, το σύστημα είναι στην κατάσταση P_{ϕ_j} , δηλαδή στην καθαρή κατάσταση που αντιστοιχεί σε ένα ιδιοδιάνυσμα του A όταν $w_j = \delta_{ij}$. Σε μια τέτοια περίπτωση η ιδιοτιμή λ_i λαμβάνεται με βεβαιότητα σε μια μέτρηση του A .

Θα θεωρήσουμε τώρα την γενική περίπτωση όπου στην ιδιοτιμή λ_i αντιστοιχούν τα ιδιοδιανύσματα $\phi_i^{(1)}, \phi_i^{(2)}, \dots, \phi_i^{(r)}$. Σε αυτή την περίπτωση θα πρέπει να αντικαταστήσουμε τις προβολές P_{ϕ_i} , με τις γραμμικές συναρτήσεις P_i που αντιστοιχούν στους ιδιόχωρους της ιδιοτιμής ϕ_i :

$$P_i \psi = \sum_{k=1}^r \langle \psi, \phi_i^{(k)} \rangle \phi_i^{(k)}.$$

Τότε η πιθανότητα w_i να πάρουμε την τιμή λ_i σε μία μέτρηση του A δίνεται από τον τύπο

$$w_i = \sum_{k=1}^r \langle M\phi_i^{(k)}, \phi_i^{(k)} \rangle \quad (\text{IX.7})$$

ενώ στην περίπτωση μιας καθαρής κατάστασης P_ψ έχουμε

$$w_i = \sum_{k=1}^r |\langle \psi, \phi_i^{(k)} \rangle|^2.$$

Έστω μια συνάρτηση f και θεωρούμε τα παρατηρήσιμα στοιχεία $A, f(A)$. Τα $A, f(A)$ αντιμετατίθενται και έχουν κοινά ιδιοδιανύσματα:

$$A\phi_i^{(k)} = \lambda_i \phi_i^{(k)}, \quad f(A)\phi_i^{(k)} = f(\lambda_i)\phi_i^{(k)}.$$

Η ποσότητα

$$\sum_{k=1}^r \langle M\phi_i^{(k)}, \phi_i^{(k)} \rangle$$

εκφράζει την πιθανότητα να πάρουμε την τιμή λ_i σε μια μέτρηση όπως και την πιθανότητα να πάρουμε την τιμή $f(\lambda_i)$ για το παρατηρήσιμο $f(A)$. Συνεπώς για $E = (a, b) \subset \mathbb{R}$

$$\omega_{f(A)}(E) = \omega_A(f^{-1}(E)).$$

Συνοψίζουμε τα στοιχεία του μοντέλου που μελετήσαμε και την φυσική σημασία τους:

1. Ένα $A \in \mathfrak{A}$ σε μία κατάσταση ω που αντιστοιχεί στον πίνακα πυκνότητας M έχει μέση τιμή

$$\langle A | \omega \rangle = \text{tr}(MA)$$

και συνάρτηση κατανομής

$$\omega_A(\lambda) = \text{tr}(MP_A(\lambda)).$$

Για μία καθαρή κατάσταση $M = P_\psi$ που αντιστοιχεί στο διάνυσμα $\psi \in \mathbb{C}^n$, $\|\psi\| = 1$ έχουμε

$$\langle A | \omega \rangle = \langle A\psi, \psi \rangle,$$

$$\omega_A(\lambda) = \langle P_A(\lambda)\psi, \psi \rangle.$$

2. Το σύνολο των ιδιοτιμών του $A \in \mathfrak{A}$ ταυτίζεται με το σύνολο των πιθανών αποτελεσμάτων μέτρησης του A .
3. Στην μέτρηση του $A \in \mathfrak{A}$, η πιθανότητα w_i να μετρήσουμε μία τιμή ίση με την ιδιοτιμή λ_i δίνεται από την (IX.7). Η πιθανότητα να μετρήσουμε μια τιμή που δεν είναι ιδιοτιμή είναι μηδενική.
4. Τα ιδιοδιανύσματα του $A \in \mathfrak{A}$ καθορίζουν τις καθарές καταστάσεις στις οποίες το A παίρνει μία τιμή ίση με την αντίστοιχη ιδιοτιμή με βεβαιότητα.

Το πεπερασμένης διάστασης αυτό μοντέλο ικανοποιεί αρκετές από τις ανάγκες που προέρχονται από την φυσική, όπως παρατηρήσιμα μεγέθη που δεν είναι ταυτόχρονα μετρήσιμα ή τις σχέσεις αβεβαιότητας. Μπορεί να περιγράψει ικανοποιητικά φυσικά φαινόμενα που έχουν ένα πεπερασμένο σύνολο δυνατών τιμών. Στην πλήρη γενικότητα, ο πεπερασμένης διάστασης χώρος V πρέπει να αντικατασταθεί από έναν απειροδιάστατο χώρο Hilbert σαν αυτούς που θα συναντήσετε στο μάθημα της συναρτησιακής ανάλυσης, ενώ ο χώρος των παρατηρήσιμων μεγεθών θα αποτελείται από τον χώρο των αυτοσυζυγών (μη-φραγμένων) γραμμικών συναρτήσεων στον χώρο αυτό. Η θεωρία αυτή, για τις ανάγκες της κβαντικής μηχανικής αναπτύχθηκε από τον von Neumann.

Ο **John von Neumann** (28 Δεκεμβρίου 1903 – 8 Φεβρουαρίου 1957) ήταν Ούγγρος Μαθηματικός, Φυσικός και μηχανικός. Τα επιστημονικά του ενδιαφέροντα και οι συνεισφορές του στην επιστήμη ήταν ευρύτατα. Ήταν ένα παιδί θαύμα που σε ηλικία 6 ετών μπορούσε να κάνει διαιρέσεις 8ψήφιων αριθμών στο κεφάλι του ενώ σε ηλικία 8 ετών γνώριζε απειροστικό λογισμό. Ήταν πρωτοπόρος στην ανάπτυξη των μαθηματικών εργαλείων στην συναρτησιακή ανάλυση για την αυστηρή θεμελίωση της κβαντομηχανικής. Κατά την διάρκεια του 2ου παγκόσμιου πολέμου εργάστηκε στο Manhattan Project για την κατασκευή πυρηνικών όπλων. Έλαβε σειρά βραβείων και τιμών και ένας κρατήρας στην σελήνη έχει το όνομα του.



IX.3 Το θεώρημα Perron-Frobenius και ο αλγόριθμος της Google

Σε αυτή την παράγραφο θα μιλήσουμε για ένα θεώρημα που αποδείχθηκε το 1907 από τον O. Perron για θετικούς πίνακες και επεκτάθηκε μη αρνητικούς πίνακες το 1912 από τον F.G. Frobenius. Το θεώρημα αυτό έχει πάρα πολλές εφαρμογές καταλήγουμε παρουσιάζοντας μια εφαρμογή του στον pagerank αλγόριθμο της Google η οποία άλλαξε τον κόσμο και έκανε δισεκατομμυριούχους τους επινοητές της μεθόδου.

Υπάρχουν πάρα πολλές αποδείξεις των θεωρημάτων αυτών, πολλές χρησιμοποιούν προχωρημένες τεχνικές από την τοπολογία όπως το θεώρημα σταθερού σημείου. Δίνουμε μια ολοκληρωμένη στοιχειώδη απόδειξη βασισμένη αφενός σε σημειώσεις του Vladimir Drinfeld, όπως αυτές επεξεργάστηκαν από τον S. Khim, [13] όπως και σε μία σύντομη απόδειξη του θεωρήματος του Perron [3] από την H. Cairns.

Ορισμός IX.3.1. Για πίνακες $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{n,m}$ θα γράφουμε $A \geq B$ (αντίστοιχα $A > B$) αν $a_{ij} \geq b_{ij}$ (αντίστοιχα $a_{ij} > b_{ij}$) για κάθε $1 \leq i \leq n, 1 \leq j \leq m$. Ένας πίνακας θα λέγεται θετικός αν $A > 0$ και θα λέγεται μη αρνητικός αν και μόνο αν $A \geq 0$.

Ορισμός IX.3.2. Έστω $A \in \mathbb{C}^{n,n}$, η φασματική του ακτίνα $\rho(A)$ ορίζεται ως

$$\rho(A) = \max\{|\lambda_1|, \dots, |\lambda_n|\},$$

όπου $\lambda_1, \dots, \lambda_n$ είναι η ιδιοτιμές του A .

Πρόταση IX.3.3. Ένας πίνακας A έχει φασματική ακτίνα $\rho(A) < 1$ αν και μόνο αν $\lim_{n \rightarrow \infty} A^n = 0$.

Αν $\rho(A) > 1$ τότε για κάθε νόρμα πινάκων $\|\cdot\|$, έχουμε $\lim_{n \rightarrow \infty} \|A^n\| = \infty$.

Απόδειξη. Υποθέτουμε ότι $A^n \rightarrow 0$. Έστω λ μια ιδιοτιμή του A με ιδιοδιάνυσμα v . Έχουμε ότι $A^n v = \lambda^n v$, συνεπώς

$$0 = \left(\lim_{n \rightarrow \infty} A^n \right) v = \lim_{n \rightarrow \infty} (A^n v) = \lim_{n \rightarrow \infty} \lambda^n v.$$

Αφού $v \neq 0$ θα πρέπει $\lim_{n \rightarrow \infty} \lambda^n = 0$ συνεπώς $|\lambda| < 1$. Αυτό πρέπει να είναι σωστό για κάθε ιδιοτιμή άρα $\rho(A) < 1$.

Υποθέτουμε τώρα ότι $\rho(A) < 1$ και θεωρούμε αντιστρέψιμο πίνακα Q ώστε $Q^{-1}AQ = J$, όπου ο J είναι η κανονική μορφή Jordan του πίνακα A . Αρκεί να αποδείξουμε ότι για κάθε Jordan block $J(\lambda, k)$ με $|\lambda| < 1$ ισχύει ότι $\lim_{n \rightarrow \infty} J(\lambda)^n = 0$. Υπολογίζουμε ότι

$$J(\lambda, k)^n = \begin{pmatrix} \lambda^n & \binom{n}{1} \lambda^{n-1} & \binom{n}{2} \lambda^{n-2} & \dots & \binom{n}{k-1} \lambda^{n-k+1} \\ 0 & \lambda^n & \binom{n}{1} \lambda^{n-1} & \ddots & \binom{n}{k-2} \lambda^{n-k+2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda^n & \binom{n}{1} \lambda^{n-1} \\ 0 & \dots & \dots & 0 & \lambda^n \end{pmatrix}$$

από όπου προκύπτει ότι $\lim_{n \rightarrow \infty} J(\lambda, k) = 0$.

Αν $\rho(A) > 1$ τότε υπάρχει ένα Jordan block στην ανάλυση του A το οποίο τείνει στο ∞ , οπότε προκύπτει το ζητούμενο στην περίπτωση αυτή. $\square$

Θεώρημα ΙΧ.3.4 (Τύπος του Gelfand). Για κάθε νόρμα πινάκων ισχύει ότι

$$\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|^{1/n}.$$

Απόδειξη. Για κάθε $\epsilon > 0$ θεωρούμε τους πίνακες

$$A_{\pm} = \frac{1}{\rho(A) \pm \epsilon} A.$$

Είναι σαφές ότι

$$\rho(A) = \frac{\rho(A)}{\rho(A) \pm \epsilon}, \text{ και } \rho(A_{+}) < 1 < \rho(A_{-}).$$

Συνεπώς, $\lim_{n \rightarrow \infty} A_{+}^n = \mathbf{0}$, δηλαδή υπάρχει n_0^{+} , ώστε για $n > n_0^{+}$ να ισχύει

$$\|A_{+}^n\| < 1.$$

Συνεπώς,

$$\|A^n\|^{1/n} < \rho(A) + \epsilon.$$

Ομοίως, έχουμε ότι η $\|A_{-}^n\|$ δεν είναι φραγμένη, οπότε υπάρχει ένα n_0^{-} , ώστε για $n > n_0^{-}$ να έχουμε

$$\|A_{-}^n\| > 1$$

από όπου προκύπτει ότι

$$\|A^n\|^{1/n} > \rho(A) - \epsilon.$$

Συνεπώς για $n_0 = \max\{n_0^{+}, n_0^{-}\}$ έχουμε ότι

Για κάθε $\epsilon > 0$, υπάρχει $n_0 \in \mathbb{N}$, ώστε $n > n_0$ να ισχύει $\rho(A) - \epsilon < \|A^n\|^{1/n} < \rho(A) + \epsilon$

δηλαδή

$$\lim_{n \rightarrow \infty} \|A^n\|^{1/n} = \rho(A).$$

□

Ο **Israel Gelfand** (2 Σεπτεμβρίου 1913 – 5 Οκτωβρίου 2009) ήταν Ρώσος Μαθηματικός, ο οποίος είχε σημαντική συνεισφορά στην θεωρία ομάδων, την θεωρία αναπαραστάσεων και την συναρτησιακή ανάλυση. Το 1989 μεταναστεύσε στις Ηνωμένες πολιτείες. Το σεμινάριο του στο Moscow State University από το 1943 μέχρι το 1989 (όπου και συνεχίστηκε στο πανεπιστήμιο Rutgers) ήταν ονομαστό, κάλυπτε ένα ευρύ φάσμα αντικειμένων και συνέβαλε στην εκπαίδευση πολλών μαθηματικών. Είχε πλήθος διάσημων μαθητών όπως οι Endre Szemerédi, Alexandre Kirillov, Edward Frenkel, Joseph Bernstein, David Kazhdan ενώ άφησε το όνομα του σε σειρά σημαντικών θεωρημάτων όπως τα Gelfand–Mazur και Gelfand–Naimark. Έλαβε σειρά βραβείων όπως η τάξη του Λένιν και το πρώτο βραβείο Wolf.



Θεώρημα IX.3.5. Έστω ένας πίνακας $B = (b_{ij}) \in \mathbb{R}^{n,n}$ ο οποίος έχει μη αρνητικά στοιχεία $b_{ij} \geq 0$. Έστω $\lambda_1, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα B .

Τότε ισχύουν τα παρακάτω:

1. Ο πίνακας B έχει μια μη-αρνητική πραγματική ιδιοτιμή. Η μεγαλύτερη πραγματική ιδιοτιμή $\lambda(B)$ ικανοποιεί $\lambda(B) \geq |\lambda_i|$. Στην περίπτωση που $b_{ij} > 0$ ισχύει $\lambda(B) > |\lambda_i|$.
2. Στην περίπτωση που $b_{ij} > 0$ η αλγεβρική και γεωμετρική πολλαπλότητα της $\lambda(B)$ είναι ίση με 1 και το αντίστοιχο ιδιοδιάνυσμα μπορεί να κανονικοποιηθεί ώστε να έχει γνήσια θετικά στοιχεία.
3. Αν ο B έχει ένα ιδιοδιάνυσμα v με γνήσια θετικά στοιχεία, τότε η αντίστοιχη ιδιοτιμή είναι η $\lambda(B)$.

Θα αποδείξουμε πρώτα το [3](#). Υποθέτουμε ότι το v έχει ένα ιδιοδιάνυσμα $v = (v_1, \dots, v_n)^t$ ώστε $v_i > 0$ και έστω λ_v η αντίστοιχη ιδιοτιμή ώστε να ισχύει $Bv = \lambda_v v$. Παρατηρούμε ότι

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & v_n \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = D(v) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Έχουμε

$$BD(v) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \lambda_D(v) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \Rightarrow D(v)^{-1}BD(v) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \lambda_v \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Ο πίνακας $D(v)^{-1} = \text{diag}(v_1^{-1}, \dots, v_n^{-1})$ έχει θετικά στοιχεία και συνεπώς ο πίνακας $D(v)^{-1}BD(v)$ έχει μη αρνητικά στοιχεία, ίδιες ιδιοτιμές με τον B αφού είναι όμοιοι. Συνεπώς, χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι ο μη-αρνητικός πίνακας B έχει ως ιδιοδιάνυσμα το $v = (1, \dots, 1)^t$.

Η επιλογή αυτή δίνει ότι $\lambda_v = \sum_{j=1}^n b_{ij}$ για κάθε $1 \leq i \leq n$. Συνεπώς, $\lambda_v \geq 0$ και εκτός αν $B = 0$ έχουμε επιπλέον ότι $\lambda_v > 0$.

Εφοδιάζουμε το $\mathbb{C}^n$ με την $\|\cdot\|_\infty$ νόρμα την οποία θα την συμβολίζουμε με $\|\cdot\|$, δηλαδή για

$$z = (z_1, \dots, z_n)^t, \|z\| = \max_{i=1, \dots, n} |z_i|.$$

Για κάθε $z \in \mathbb{C}^n$ η i συντεταγμένη του διανύσματος Bz είναι το $b_{i1}z_1 + b_{i2}z_2 + \cdots + b_{in}z_n$, και έχουμε

$$|b_{i1}z_1 + b_{i2}z_2 + \cdots + b_{in}z_n| \leq |b_{i1}||z_1| + |b_{i2}||z_2| + \cdots + |b_{in}||z_n| \leq \sum_{j=1}^n b_{ij} \max_{i=1, \dots, n} |z_i| = \lambda_v \|z\|. \quad (\text{IX.8})$$

Συνεπώς

$$\|Bz\| \leq \lambda_v \|z\|.$$

Άρα αν z' είναι ένα ιδιοδιάνυσμα με ιδιοτιμή λ' τότε

$$\|Bz'\| = |\lambda'| \cdot \|z'\| \leq \lambda_v \|z'\|.$$

Συνεπώς $|\lambda'| \leq \lambda_v$ και $\lambda_v = \lambda(B)$.

Παρατήρηση IX.3.6. Αν υποθέσουμε ότι $b_{ij} > 0$. Τότε $\|Bz\| < \lambda_v \|z\|$, εκτός αν $z_1 = z_2 = \cdots = z_n$, το οποίο είναι ισοδύναμο με το ότι $z \in ((1, \dots, 1)^t)$. Αυτό σημαίνει ότι ο ιδιόχωρος της ιδιοτιμής λ_v είναι μονοδιάστατος. Πράγματι, αν $z_i \neq z_j$ για κάποια $1 \leq i, j \leq n$ τότε η ανισότητα [\(IX.8\)](#) θα είναι γνήσια. Αυτό σημαίνει ότι $|\lambda| < \lambda_v$ για κάθε ιδιοτιμή του B .

Παρατήρηση ΙΧ.3.7. Θα αποδείξουμε τώρα η αλγεβρική πολλαπλότητα του λ_v είναι επίσης ίση με 1, δηλαδή το λ_v είναι απλή ρίζα του χαρακτηριστικού πολυωνύμου. Έστω ότι το λ_v εμφανίζεται ως πολλαπλή ρίζα του χαρακτηριστικού πολυωνύμου. Αυτό σημαίνει ότι η κανονική μορφή Jordan του πίνακα B έχει ένα μπλοκ της μορφής:

$$J(\lambda_v, r) = \begin{pmatrix} \lambda_v & 1 & 0 & \cdots & 0 \\ 0 & \lambda_v & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda_v & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_v \end{pmatrix}$$

με $r \geq 2$. Το Jordan μπλοκ που αντιστοιχεί στην ιδιοτιμή λ_v είναι μοναδικό αφού έχουμε ήδη δείξει ότι η γεωμετρική πολλαπλότητα της ιδιοτιμής λ_v είναι 1.

Έστω Q ο πίνακας ώστε $Q^{-1}BQ$ να είναι η κανονική μορφή Jordan. Ισχυριζόμαστε τώρα ότι στους πίνακες $(\lambda_v^{-1}Q^{-1}BQ)^n$ υπάρχουν στοιχεία τα οποία συγκλίνουν στο ∞ όταν $n \rightarrow \infty$.

Υπολογίζουμε

$$\frac{1}{\lambda_v} J(\lambda_v, r) = \frac{1}{\lambda_v} \begin{pmatrix} \lambda_v & 1 & 0 & \cdots & 0 \\ 0 & \lambda_v & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda_v & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_v \end{pmatrix} = \mathbb{I}_r + \begin{pmatrix} 0 & \frac{1}{\lambda_v} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{\lambda_v} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & \frac{1}{\lambda_v} \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} \mathbb{I}_r + A$$

Υπολογίζουμε ότι

$$\begin{aligned} (\mathbb{I}_r + A)^n &= \sum_{k=0}^n \binom{n}{k} A^k = \mathbb{I}_r + nA + \binom{n}{2} A^2 + \cdots \\ &= \begin{pmatrix} 1 & \frac{n}{\lambda_v} & 0 & \cdots & 0 \\ 0 & 1 & \frac{n}{\lambda_v} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 & \frac{n}{\lambda_v} \\ 0 & \cdots & \cdots & 0 & \frac{n}{\lambda_v} \end{pmatrix} \end{aligned}$$

Το οποίο αποδεικνύει ότι ο πίνακας $(\lambda_v^{-1}Q^{-1}BQ)^n$ έχει στοιχεία τα οποία συγκλίνουν στο ∞ . Για το παραπάνω χρησιμοποιούμε ότι ο πολλαπλασιασμός μπλοκ διαγώνιων πινάκων ανάγεται στον υπολογισμό του κάθε μπλοκ.

Το αποτέλεσμα για τον πίνακα B^n είναι επίσης σαφές αφού $B^n = Q(Q^{-1}BQ)^nQ^{-1}$ και η ιδιοτιμή λ_v είναι πεπερασμένη. Όμως καταλλήγουμε σε άτοπο αφού

$$\|Bz\| \leq \lambda_v \|z\| \Leftrightarrow \frac{1}{\lambda_v} \|Bz\| = \left\| \frac{1}{\lambda_v} Bz \right\| \leq \|z\|$$

και συνεπώς

$$\left\| \left(\frac{1}{\lambda_v} B \right)^n z \right\| = \left\| \frac{1}{\lambda_v} \left(\frac{1}{\lambda_v} \right)^{n-1} B^n z \right\| \leq \|z\| \quad (\text{IX.9})$$

για κάθε $\|z\|$, αφού ο πίνακας $\left(\frac{1}{\lambda_v} \right)^{n-1} B^n$ είναι ένας πίνακας με στοιχεία γνήσια μεγαλύτερα του μηδενός. Παρατηρούμε ότι πίνακες A με την ιδιότητα $\|Az\| \leq \|z\|$ έχουν την ιδιότητα $\|az\| \leq \|z\|$ όπου a είναι μια τυχαία γραμμή του πίνακα A. Τότε τα στοιχεία του A είναι φραγμένα αφού $|a_{ij}| \leq 2 \frac{\|z\|}{|z_j|}$. Συνεπώς τα στοιχεία του $\left(\frac{1}{\lambda_v} B \right)^n$ είναι φραγμένα και δεν μπορούν να τείνουν στο άπειρο. Άρα η αλγεβρική πολλαπλότητα της ιδιοτιμής λ_v είναι ίση με 1.

Θα ανάγουμε την απόδειξη των 1 και 2 στην περίπτωση που ο πίνακας B έχει στοιχεία $b_{ij} > 0$. Έστω ένας πίνακας $B = (b_{ij})$ με $b_{ij} \geq 0$. Θεωρούμε τον πίνακα B_r ο οποίος είναι ο πίνακας B στον οποίο τα μηδενικά στοιχεία έχουν αντικατασταθεί από τα στοιχεία $1/r$, με $r \in \mathbb{R}$, $r > 0$. Θα δείξουμε ότι

- (i) οι ιδιοτιμές του πίνακα B_r συγκλίνουν στις ιδιοτιμές του B όταν $r \rightarrow \infty$.
- (ii) θα αποδείξουμε ότι υπάρχει ένα ιδιοδιάνυσμα του πίνακα B_r με γνήσια θετικά στοιχεία και συνεπώς υπάρχει μια θετική ιδιοτιμή του B_r από την απόδειξη του 3.
- (iii) Θα αποδείξουμε ότι η ιδιοτιμή αυτή είναι η $\lambda(B)$ και ικανοποιεί τις ιδιότητες 1 και 2.

Οι ιδιοτιμές ενός πίνακα είναι οι ρίζες του χαρακτηριστικού του πολυωνύμου. Θα δείξουμε ότι οι ιδιοτιμές των B_r συγκλίνουν στις ιδιοτιμές του B .

Λήμμα IX.3.8. Έστω $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$, $a_i \in \mathbb{C}$, και $|a_1|, \dots, |a_n| < M$. Αν z είναι μια ρίζα του f , τότε $|z| < 1 + nM$.

Απόδειξη. Υποθέτουμε ότι $|z| \geq 1 + nM$. Αν $f(z) = 0$, τότε έχουμε

$$|z^n| = |-a_{n-1}z^{n-1} - a_{n-2}z^{n-2} - \dots - a_0|$$

οπότε έχουμε

$$\begin{aligned} 1 &= |a_{n-1}z^{n-1}z^n + a_{n-2}z^{n-2}z^n + \dots + a_0z^n| \\ &\leq \left| \frac{a_{n-1}}{z} \right| + \left| \frac{a_{n-2}}{z^2} \right| + \dots + \left| \frac{a_0}{z^n} \right| \\ &< \left| \frac{M}{z} \right| + \left| \frac{M}{z^2} \right| + \dots + \left| \frac{M}{z^n} \right| \\ &\leq \left| \frac{M}{1+nM} \right| + \left| \frac{M}{(1+nM)^2} \right| + \dots + \left| \frac{M}{(1+nM)^n} \right| \\ &\leq n \left| \frac{M}{1+nM} \right| < 1 \end{aligned}$$

το οποίο είναι άτοπο. □

Έστω $\lambda_1^{(r)}, \dots, \lambda_n^{(r)}$ οι ρίζες του B_r . Η ακολουθία $\{(\lambda_1^{(r)}, \dots, \lambda_n^{(r)})\}_{r=1}^\infty$ είναι φραγμένη, συνεπώς έχει συγκλίνουσα υποακολουθία $\{(\lambda_1^{(r_j)}, \dots, \lambda_n^{(r_j)})\}_{j=1}^\infty$. Επίσης $B_{r_j} \rightarrow B$, άρα $\text{ch}_{B_{r_j}} \rightarrow \text{ch}_B$. Αν θέσουμε $\lambda_k = \lim_{j \rightarrow \infty} \lambda_k^{(r_j)}$ τότε έχουμε

$$\text{ch}_B(x) = \prod_{k=1}^n (\lambda_k - x)$$

αφού

$$\text{ch}_{B_{r_j}}(x) = \prod_{k=1}^n (\lambda_k^{(r_j)} - x).$$

Δηλαδή τα όρια λ_k είναι οι ιδιοτιμές του B . Αυτό αποδεικνύει το (i).

Λήμμα IX.3.9. Έστω $A \in \mathbb{R}^{n,n}$ ένας θετικός πίνακας. Αν θ, ψ είναι δύο μη ίσες στήλες του πίνακα με $\theta \geq \psi$, τότε $A\theta > A\psi$. Υπάρχει $\epsilon > 0$ με $\theta > (1 + \epsilon)A\psi$.

Απόδειξη. Το διάνυσμα $A(\theta - \psi) = \sum_{j=1}^n a_{ij}(\theta_j - \psi_j)$, οπότε

$$A(\theta - \psi)_i \geq \sum_{j=1}^n \min_{ij} \{a_{ij}\} (\theta_j - \psi_j) = \min_{ij} \{a_{ij}\} \sum_{j=1}^n (\theta_j - \psi_j) > 0.$$

Συνεπώς, $(A\theta)_i > (A\psi)_i$ και εξορισμού έχουμε $A\theta > A\psi$. Αν διαταράξουμε το διάνυσμα $A(\theta - \psi)$ με μικρή διαταραχή εξακολουθεί να παραμένει θετικό, δηλαδή υπάρχει $\epsilon > 0$ με $A(\theta - \psi) - \epsilon A\psi > 0$, δηλαδή $A\theta > (1 + \epsilon)A\psi$. $\square$

Υπάρχει μία ιδιοτιμή λ με $|\lambda| = \rho(A)$ και έστω ψ το αντίστοιχο ιδιοδιάνυσμα. Θεωρούμε το διάνυσμα $\Psi = (\Psi_1, \dots, \Psi_n)^t$ με $\Psi_j = |\psi_j|$, όπου $\psi = (\psi_1, \dots, \psi_n)^t$. Έχουμε

$$(A\Psi)_i = \sum_{j=1}^n a_{ij}|\psi_j| \geq \left| \sum_{j=1}^n a_{ij}\psi_j \right| = |\lambda\psi| = \rho(A)\Psi_i,$$

συνεπώς $A\Psi \geq \rho(A)\Psi$. Αν η παραπάνω ανισότητα είναι γνήσια τότε από το λήμμα IX.3.9 έχουμε ότι $A^2\Psi \geq \rho(A)A\Psi$ και υπάρχει $\epsilon > 0$ με $A^2\Psi \geq (1 + \epsilon)\rho(A)A\Psi$. Ο πίνακας A είναι μη-αρνητικός και το ίδιο είναι και οι δυνάμεις του A^n από όπου έχουμε

$$A^{n+2}\Psi \geq (1 + \epsilon)\rho(A)A^{n+1}\Psi.$$

Το παραπάνω είναι σωστό για κάθε $n \in \mathbb{N}$, οπότε

$$\begin{aligned} A^{n+1}\Psi &\geq (1 + \epsilon)\rho(A)A^n\Psi \\ &\geq ((1 + \epsilon)\rho(A))^2 A^{n-1}\Psi \\ &\dots \\ &\geq ((1 + \epsilon)\rho(A))^n A^{n-1}\Psi. \end{aligned}$$

Από τον τύπο του Gelfand έχουμε ότι

$$\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|^{1/n} \geq (1 + \epsilon)\rho(A).$$

Το τελευταίο είναι άτοπο, συνεπώς $A\Psi = \rho(A)\Psi$. Δηλαδή το Ψ είναι ιδιοδιάνυσμα με ιδιοτιμή $\rho(A)$. Επίσης είναι θετικό αφού το $A\Psi = \rho(A)\Psi$ είναι θετικό. Αυτό αποδεικνύει το (ii).

Έχουμε δείξει ότι για κάθε B_r θετικό πίνακα, υπάρχει ένα θετικό ιδιοδιάνυσμα με μία θετική μέγιστη ιδιοτιμή ίση με $\rho(B_r)$, γνήσια μεγαλύτερη από όλες τις ιδιοτιμές του B_r . Αναδιατάσσουμε τις ιδιοτιμές του B_r ώστε $\lambda_1^{(r)} = \rho(B_r)$. Η ακολουθία $\{((\lambda_1^{(r_1)}, \dots, \lambda_n^{(r_1)}))\}_{j=1}^\infty$ συγκλίνει στην $(\lambda_1, \dots, \lambda_n)$.

Καταλήγουμε στο ότι $\lambda_1^{(r_j)} > 0$ συνεπώς $\lambda_1 \in \mathbb{R}$ και $\lambda_1 \in \mathbb{R}$. Επίσης

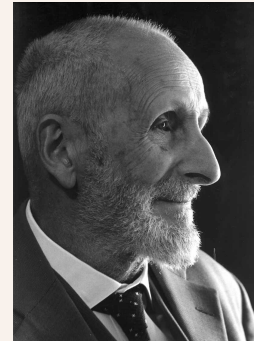
$$\lambda_1^{(r_j)} > |\lambda_k^{(r_j)}| \text{ για κάθε } 2 \leq k \leq n$$

από όπου έχουμε ότι $\lambda_1 \geq |\lambda_k|$ για κάθε $2 \leq k \leq n$. Αυτό ολοκληρώνει την απόδειξη του θεωρήματος IX.3.5.

Θεώρημα IX.3.10. Ένας μη-αρνητικός πίνακας $A \leq 0$ για τον οποίο ισχύει $A^m > 0$ ικανοποιεί ότι $\rho(A)$ είναι ιδιοτιμή και ότι έχει θετικό ιδιοδιάνυσμα. Πίνακες αυτής της μορφής ονομάζονται πρωταρχικοί (primitive). Η ιδιοτιμή $\rho(A)$ είναι η μοναδική ιδιοτιμή λ με $|\lambda| = \rho(A)$ και η $\rho(\lambda)$ έχει αλγεβρική και γεωμετρική πολλαπλότητα ίση με 1.

Απόδειξη. Έστω $\lambda_1, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα A και έστω λ_1 η ιδιοτιμή με την μεγαλύτερη απόλυτη τιμή. Οι ιδιοτιμές του A^m είναι οι $\lambda_1^m, \dots, \lambda_n^m$ και έχουμε ότι $\lambda_1^m > 0$ και έχει θετικό ιδιοδιάνυσμα ψ . Όλες οι άλλες ιδιοτιμές του A^m είναι μικρότερες από την λ_1^m σε απόλυτη τιμή. Όμως αν $|\lambda_1|^m > |\lambda_j|^m$ για $j \geq 2$ τότε και $\lambda_1 = |\lambda_1| > |\lambda_j|$ για $j \geq 2$. Η λ_1 έχει και αυτή ιδιοδιάνυσμα ψ , συνεπώς η αλγεβρική και γεωμετρική πολλαπλότητα της $\lambda_1 = \rho(A)$ είναι ίση με 1. $\square$

Ο Oskar Perron (7 Μαΐου 1880 - 22 Φεβρουαρίου 1975) ήταν Γερμανός Μαθηματικός που εργάστηκε στον τομέα των διαφορικών και μερικών διαφορικών εξισώσεων αλλά και στην θεωρία των συνεχών κλασμάτων την Γεωμετρία και την Θεωρία Αριθμών. Φοίτησε στα πανεπιστήμια του Tübingen και Göttingen University όπου και συνεργάστηκε με τον David Hilbert. Εργάστηκε στα πανεπιστήμια της Χαιδελβέργης και του Μονάχου. Η συνεισφορά του στην Ανάλυση περιλαμβάνει το ολοκλήρωμα Perron και είναι ευρύτατα γνωστός για το θεώρημα Perron-Frobenius, όπως και για το παράδοξο που φέρει το όνομα του.



Ο Ferdinand Georg Frobenius (26 Οκτωβρίου 1849 - 3 Αυγούστου 1917) ήταν Γερμανός Μαθηματικός με σημαντική συνεισφορά στην θεωρία των Ελλειπτικών συναρτήσεων, στις διαφορικές εξισώσεις, την θεωρία Αριθμών, την Θεωρία Ομάδων και την Θεωρία Αναπαραστάσεων.

Η θεωρία των αλγεβρών Frobenius πέρα από την σημασία της για την θεωρία αναπαραστάσεων έχει εφαρμογές στις τοπολογικές κβαντικές θεωρίες πεδίου (TQFT) ενώ ο αυτομορφισμός που φέρει το όνομα του είναι ένα σημαντικό εργαλείο στην Αλγεβρική Θεωρία Αριθμών με εφαρμογές στην επιστήμη υπολογιστών σχετικές με την καταμέτρηση ρητών σημείων αλγεβρικών πολλαπλοτήτων πάνω από πεπερασμένα σώματα. Το θεώρημα Perron-Frobenius έχει σημαντικές εφαρμογές στην θεωρία των δυναμικών συστημάτων και στα εφαρμοσμένα και οικονομικά Μαθηματικά.



IX.3.1 Στοχαστικοί πίνακες

Ορισμός IX.3.11. Ένας πίνακας θα λέγεται στοχαστικός, αν είναι μη αρνητικός και το άθροισμα κατά μήκος των στηλών είναι ίσο με 1.

Παράδειγμα IX.3.12. Ο πίνακας A είναι θετικός στοχαστικός ενώ ο πίνακας $\mathbb{I}_3$ είναι στοχαστικός αλλά όχι θετικός.

$$A = \begin{pmatrix} 0.3 & 0.2 & 0.6 \\ 0.1 & 0.5 & 0.2 \\ 0.6 & 0.3 & 0.2 \end{pmatrix} \quad \mathbb{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Πρόταση ΙΧ.3.13. Έστω A ένας στοχαστικός πίνακας. Τότε το 1 είναι ιδιοτιμή του A . Κάθε άλλη πραγματική ή μιγαδική ιδιοτιμή λ του A ικανοποιεί την σχέση $|\lambda| \leq 1$. $|\lambda|$

Απόδειξη. Αφού ο A είναι στοχαστικός έχουμε ότι

$$A^t \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

δηλαδή το $(1, \dots, 1)^t$ είναι ιδιοδιάνυσμα του A^t με ιδιοτιμή 1. Αφού ο A και A^t έχουν το ίδιο χαρακτηριστικό πολυώνυμο και συνεπώς τις ίδιες ιδιοτιμές έχουμε ότι το 1 είναι ιδιοτιμή του A .

Έστω λ μια άλλη ιδιοτιμή του A άρα και του A^t . Έστω $x = (x_1, \dots, x_n)^t$ ένα ιδιοδιάνυσμα του A^t της ιδιοτιμής λ , δηλαδή $A^t x = \lambda x$. Συγκρίνοντας την j συντεταγμένη παίρνουμε

$$\lambda x_j = \sum_{i=1}^n a_{ij} x_i.$$

Θεωρούμε το j ώστε $|x_i| \leq |x_j|$ για κάθε $1 \leq i \leq n$. Έχουμε

$$|\lambda| \cdot |x_j| = \left| \sum_{i=1}^n a_{ij} x_i \right| \leq \sum_{i=1}^n a_{ij} |x_i| = |x_j|$$

όπου η τελευταία σχέση ισχύει γιατί ο A είναι στοχαστικός και συνεπώς $\sum_{i=1}^n a_{ij} = 1$. Καταλήγουμε στο ότι $|\lambda| \leq 1$. $\square$

Παρατήρηση ΙΧ.3.14. Το θεώρημα [IX.3.5](#) δίνει ότι αν ο A είναι θετικός στοχαστικός τότε $|\lambda| < 1$ για όλες τις ιδιοτιμές που δεν είναι μηδενικές.

Παράδειγμα ΙΧ.3.15. Για τον πίνακα

$$A = \begin{pmatrix} 0.3 & 0.2 & 0.6 \\ 0.1 & 0.5 & 0.2 \\ 0.6 & 0.3 & 0.2 \end{pmatrix}$$

υπολογίζουμε ότι οι ιδιοτιμές είναι $1, 0.360555, -0.360555$.

Ορισμός ΙΧ.3.16. Μια σταθερή κατάσταση του στοχαστικού πίνακα A είναι το ιδιοδιάνυσμα της ιδιοτιμής 1 το οποίο έχει κανονικοποιηθεί ώστε τα στοιχεία του να έχουν άθροισμα 1.

Για παράδειγμα για τον πίνακα υπολογίζουμε ότι το διάνυσμα σταθερής κατάστασης είναι το

$$w = (0.390805, 0.229885, 0.37931)^t.$$

Το θεώρημα Perron-Frobenius εξασφαλίζει ότι θετικοί στοχαστικοί πίνακες έχουν ένα μοναδικό ιδιοδιάνυσμα w σταθερής κατάστασης.

Έστω ένα διάνυσμα $v_0 \in \mathbb{C}^n$ το οποίο έχει άθροισμα στοιχείων c . Υπολογίζουμε τα

$$v_1 = Av_0, v_2 = Av_1, \dots, v_t = Av_{t-1}, \dots$$

τα οποία συγκλίνουν στο cw .

Πράγματι ένας θετικός στοχαστικός πίνακας A έχει κανονική μορφή Jordan η οποία έχει ένα Jordan block $J(1, 1)$ που αντιστοιχεί στην ιδιοτιμή 1 που εμφανίζεται με πολλαπλότητα 1 και Jordan blocks $J(\lambda_i, k_i)$ στα οποία η ιδιοτιμή λ_i είναι σε απόλυτη τιμή γνήσια μικρότερη του

1. Δηλαδή για κάποιον πίνακα Q , του οποίου η πρώτη στήλη είναι το ιδιοδιάνυσμα σταθερής κατάστασης w , έχουμε

$$Q^{-1}AQ = \left(\begin{array}{c|c|c|c} 1 & 0 & \cdots & 0 \\ \hline 0 & J(\lambda_1, k_1) & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & J(\lambda_r, k_r) = J \end{array} \right) \text{ συνεπώς } \lim_{n \rightarrow \infty} A^n v_0 = Q \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array} \right) Q^{-1} v_0$$

Θεωρούμε την βάση $B = \{w, b_2, \dots, b_n\}$ στην οποία ο πίνακας A έχει την μορφή J . Ο πίνακας Q^{-1} μεταφέρει συντεταγμένες από την αρχική βάση στην βάση B , ενώ ο πίνακας Q μεταφέρει συντεταγμένες από την βάση B στην αρχική βάση. Συνεπώς $Q^{-1}v_0$ είναι οι συντεταγμένες του v_0 στην βάση B . Το

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array} \right) Q^{-1} v_0$$

είναι η προβολή του διανύσματος $Q^{-1}v_0$ στον χώρο $\langle w \rangle$. Με άλλα λόγια στην βάση B , διατηρείται μόνο η πρώτη συντεταγμένη:

$$Q^{-1}v_0 = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ τότε } \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array} \right) Q^{-1}v_0 = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = y.$$

Τελικά καταλήγουμε στον πίνακα $Qy = x_0 w$. Από την άλλη αφού ο πίνακας A είναι στοχαστικός το άθροισμα των συντεταγμένων c του $v_0, v_1 = Av_0, \dots, v_n = Av_{n-1}, \dots$ παραμένει σταθερό και αναγκαστικά το $x_0 = c$.

Παράδειγμα IX.3.17. Ο πίνακας αλλαγής βάσης με την πρώτη στήλη το διάνυσμα σταθερής κατάστασης είναι ο

$$Q = \begin{pmatrix} 0.390805 & 0.490799 & 0.652521 \\ 0.229885 & -0.810499 & 0.0987837 \\ 0.37931 & 0.3197 & -0.751304 \end{pmatrix}$$

ενώ για το διάνυσμα $v_0 = (x_1, x_2, x_3)^t$ υπολογίζουμε

$$Q^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + x_3 \\ 0.364048x_1 - 0.937249x_2 + 0.19295x_3 \\ 0.659781x_1 + 0.106044x_2 - 0.744044x_3 \end{pmatrix} \text{ και } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + x_3 \\ 0 \\ 0 \end{pmatrix}$$

Στον παρακάτω πίνακα λογαριάζουμε τις διαδοχικές δυνάμεις του $A^i v$, $i = 1, \dots, 14$ για $v = (1/6, 2/6, 3/6)^t$ και $v = (1/5, 2/5, 2/5)^t$.

(0.166667, 0.333333, 0.5)	(0.2, 0.4, 0.4)	(IX.10)
(0.416667, 0.283333, 0.3)	(0.38, 0.3, 0.32)	
(0.361667, 0.243333, 0.395)	(0.366, 0.252, 0.382)	
(0.394167, 0.236833, 0.369)	(0.3894, 0.239, 0.3716)	
(0.387017, 0.231633, 0.38135)	(0.38758, 0.23276, 0.37966)	
(0.391242, 0.230788, 0.37797)	(0.390622, 0.23107, 0.378308)	
(0.390312, 0.230112, 0.379576)	(0.390385, 0.230259, 0.379356)	
(0.390861, 0.230002, 0.379136)	(0.390781, 0.230039, 0.37918)	
(0.390741, 0.229915, 0.379345)	(0.39075, 0.229934, 0.379316)	
(0.390812, 0.2299, 0.379288)	(0.390802, 0.229905, 0.379293)	
(0.390796, 0.229889, 0.379315)	(0.390798, 0.229891, 0.379311)	
(0.390806, 0.229887, 0.379307)	(0.390804, 0.229888, 0.379308)	
(0.390804, 0.229886, 0.379311)	(0.390804, 0.229886, 0.37931)	
(0.390805, 0.229885, 0.37931)	(0.390805, 0.229885, 0.37931)	
$w^t = (0.390805, 0.229885, 0.37931)$	$w^t = (0.390805, 0.229885, 0.37931)$	

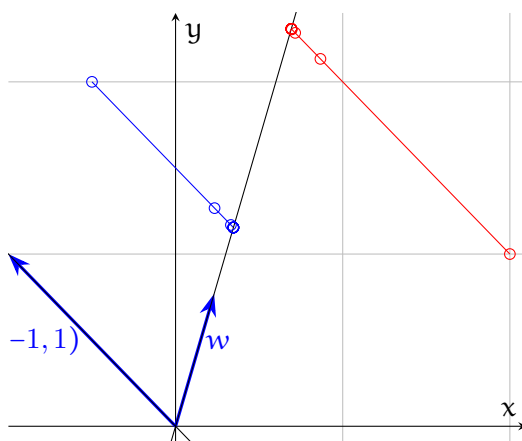
Παράδειγμα ΙΧ.3.18. Θεωρούμε τον στοχαστικό πίνακα

$$A = \begin{pmatrix} 1/3 & 1/5 \\ 2/3 & 4/5 \end{pmatrix}$$

ο οποίος έχει ιδιοτιμές $1, 2/15$. Το διάνυσμα σταθερής κατάστασης είναι το $w = (3/13, 10/13)$ και ο πίνακας αλλαγής βάσης είναι ο

$$Q = \begin{pmatrix} \frac{3}{13} & -1 \\ \frac{10}{13} & 1 \end{pmatrix}$$

Στο σχήμα **IX.2** ζωγραφίζουμε τον ιδιόχωρο της ιδιοτιμής 1 και της $2/15$ και υπολογίζουμε διαδοχικές δυνάμεις σημείων. Παρατηρούμε ότι οι διαδοχικές δυνάμεις προβάλλονται παράλληλα κατά μήκος του ιδιόχωρου $(-1, 1)^t$.



Σχήμα IX.2: Υπολογίζουμε τις δυνάμεις $A^i v$ για τα σημεία $v = (-1/2, 2)$ (μπλε χρώμα) και $v = (2, 1)$ (κόκκινο χρώμα).

Για να υπολογίσουμε το διάνυσμα σταθερής κατάστασης μπορούμε να ακολουθήσουμε τους δύο διαφορετικούς τρόπους:

Μέθοδος 7 (Υπολογισμός διανύσματος σταθερής κατάστασης). Εκτελούμε τα παρακάτω βήματα:

- 1 Υπολογίζουμε ένα ιδιοδιάνυσμα της ιδιοτιμής 1 λύνοντας την εξίσωση $(A - I_n)v = 0$.
- 2 Επιλέγουμε μία λύση v την οποία την διαιρούμε με το άθροισμα των συντεταγμένων.
- 3 Το διάνυσμα που προκύπτει είναι θετικό και οι συντεταγμένες του έχουν άθροισμα 1 και είναι το διάνυσμα κατάστασης.

Η παρακάτω μέθοδος είναι γρηγορότερη αν γίνεται με υπολογιστή.

Μέθοδος 8 (Υπολογισμός σταθερής κατάστασης). Εκτελούμε τα παρακάτω βήματα:

- 1 Επιλέγουμε ένα διάνυσμα v του οποίου τα στοιχεία αθροίζονται στο 1 .
- 2 Υπολογίζουμε τις διαδοχικές δυνάμεις $A^i v$ οι οποίες σε αριθμητική πεπερασμένη ακρί-

βειας κινητής υποδιαστολής σταθεροποιούνται στο διάνυσμα σταθερής κατάστασης.

Για ένα παράδειγμα της δεύτερης μεθόδου δείτε τους υπολογισμούς στους πίνακες της εξίσωσης (IX.10).

IX.3.2 Ο pagerank αλγόριθμος

Την δεκαετία του 1990, στα πρώτα βήματα της ανάπτυξης του internet η αναζήτηση στις υπάρχουσες ιστοσελίδες δεν ήταν αποτελεσματική. Οι υπάρχουσες μηχανές αναζήτησης όπως το Yahoo ή το AltaVista διέτρεχαν το δίκτυο καταγράφοντας τις σελίδες και τους συνδέσμους τους και στην αναζήτηση μιας λέξης οι μηχανές εμφάνιζαν πρώτα τις σελίδες στις οποίες η λέξη εμφανίζονταν τις περισσότερες φορές μέσα σε αυτή.

Οι Larry Page και Sergey Brin διδακτορικοί φοιτητές της επιστήμης υπολογιστών στο πανεπιστήμιο Stanford το 1998 και δισεκατομμυριούχοι σήμερα συνέλαβαν τον εξής αλγόριθμο ταξινόμησης ιστοσελίδων με βάση την «σημαντικότητα» κάθε σελίδας.

Κάθε ιστοσελίδα έχει μια σημαντικότητα ή rank η οποία προσδιορίζεται ως εξής:

Αν η σελίδα P συνδέεται και οδηγεί σε n το πλήθος άλλες σελίδες $Q_1, \dots, Q_n$ τότε κάθε μία από αυτές κληρονομεί το $1/n$ της σημαντικότητας της σελίδας P . Δηλαδή:

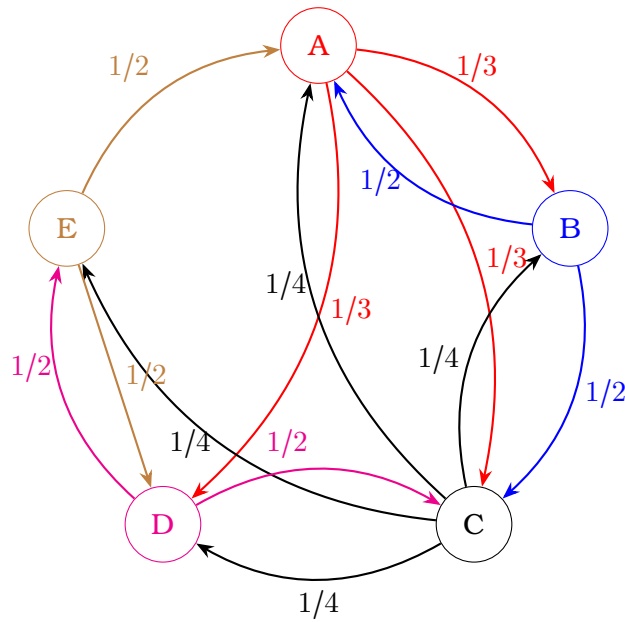
- Αν μία πολύ σημαντική σελίδα οδηγεί στην ιστοσελίδα μου (και όχι σε μερικά δισεκατομμύρια άλλες ιστοσελίδες), τότε η ιστοσελίδα μου είναι σημαντική.
- Αν ένα δισεκατομμύριο μη σημαντικές σελίδες οδηγούν στην ιστοσελίδα μου τότε αυτή είναι σημαντική.
- Αν μόνο λίγες μη σημαντικές ιστοσελίδες οδηγούν στην ιστοσελίδα μου τότε αυτή δεν είναι σημαντική.

Ορισμός IX.3.19. Θεωρούμε το διαδίκτυο ως ένα πεπερασμένο γράφημα με n -σελίδες που παίζουν τον ρόλο των κορυφών. Ο πίνακας σημαντικότητας (importance matrix) είναι ένας $n \times n$ πίνακας του οποίου το i, j -στοιχείο είναι η σημαντικότητα που η σελίδα j κληρονομεί στην σελίδα i .

Παρατήρηση IX.3.20. Παρατηρούμε ότι ο πίνακας σημαντικότητας είναι ένας στοχαστικός πίνακας υποθέτοντας ότι κάθε σελίδα περιέχει ένα σύνδεσμο προς κάποια άλλη σελίδα. Αν η σελίδα i έχει m -συνδέσμους προς άλλες σελίδες τότε η i -στήλη περιέχει το $1/m$ m το πλήθος φορές και 0 στα στοιχεία του πίνακα με τα οποία η σελίδα i δεν συνδέεται.

Παράδειγμα IX.3.21. Θεωρούμε το παρακάτω γράφημα διαδικτύου το οποίο περιέχει 5 ιστοσελίδες, τις A, B, C, D, E. Έχουμε ότι

- Η σελίδα A οδηγεί σε τρεις σελίδες και μεταφέρει $1/3$ στις σελίδες B, C, D με τις οποίες συνδέεται.
- Η σελίδα B οδηγεί σε δύο σελίδες και μεταφέρει $1/2$ στις σελίδες A, C με τις οποίες συνδέεται.
- Η σελίδα C οδηγεί σε τέσσερις σελίδες και μεταφέρει $1/4$ στις σελίδες A, B, D, E με τις οποίες συνδέεται.
- Η σελίδα D οδηγεί σε δύο σελίδες και μεταφέρει $1/2$ στις σελίδες C, E με τις οποίες συνδέεται.



- Η σελίδα **E** οδηγεί σε δύο σελίδες και μεταφέρει $1/2$ στις σελίδες A, D με τις οποίες συνδέεται.

Υπολογίζουμε τον πίνακα σημαντικότητας ως:

$$A = \begin{pmatrix} 0 & 1/2 & 1/4 & 0 & 1/2 \\ 1/3 & 0 & 1/4 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 & 0 \\ 1/3 & 0 & 1/4 & 0 & 1/2 \\ 0 & 0 & 1/4 & 1/2 & 0 \end{pmatrix} \quad (\text{IX.11})$$

Δηλαδή σε όρους πινάκων αν $v = (x_A, x_B, x_C, x_D, x_E)^t$ είναι ο πίνακας που περιέχει την rank των σελίδων A, B, C, D, E τότε

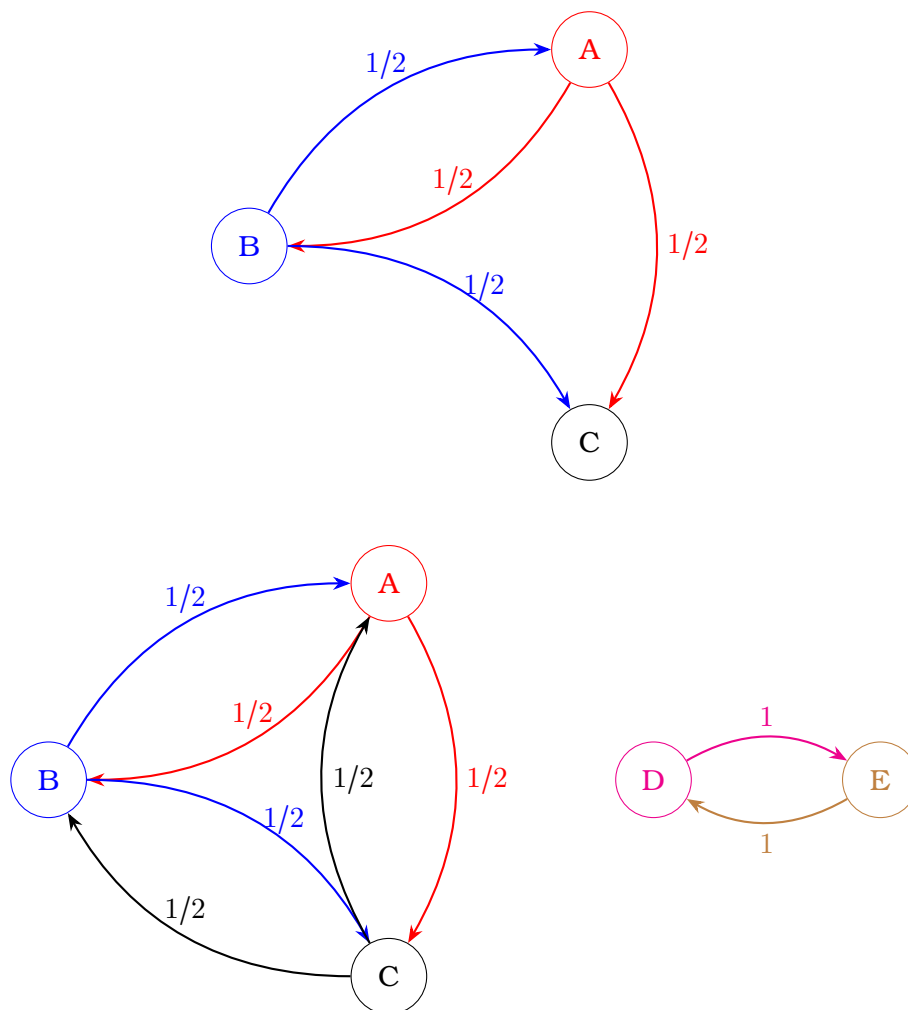
$$\begin{pmatrix} 0 & 1/2 & 1/4 & 0 & 1/2 \\ 1/3 & 0 & 1/4 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 & 0 \\ 1/3 & 0 & 1/4 & 0 & 1/2 \\ 0 & 0 & 1/4 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} x_A \\ x_B \\ x_C \\ x_D \\ x_E \end{pmatrix} = \begin{pmatrix} x_A \\ x_B \\ x_C \\ x_D \\ x_E \end{pmatrix}$$

Παρατήρηση IX.3.22. Το διάνυσμα των τάξεων είναι ιδιοδιάνυσμα της ιδιοτιμής 1 του πίνακα σημαντικότητας.

Παράδειγμα IX.3.23. Μια σελίδα χωρίς συνδέσμους Ο πίνακας σημαντικότητας του παραπάνω διαδικτύου είναι ο

$$I_1 = \begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

ο οποίος δεν είναι στοχαστικός. Η τελευταία στήλη αποτελείται από μηδενικά αφού δεν υπάρχει σύνδεσμος από την σελίδα C σε κάποια άλλη. Το χαρακτηριστικό πολυώνυμο του παραπάνω πίνακα είναι $-x^3 + x/4$ και οι ιδιοτιμές είναι οι 0, $-1/2$, $1/2$. Δεν υπάρχει ιδιοδιάνυσμα στην ιδιοτιμή 1, δηλαδή δεν υπάρχει διάνυσμα τάξεων.



Παράδειγμα IX.3.24. Μη συνεκτικό γράφημα Ο πίνακας σημαντικότητας είναι ο

$$I_2 = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Οι ιδιοτιμές του παραπάνω πίνακα είναι οι $\pm 1, \pm 1/2$ ενώ το χαρακτηριστικό πολυώνυμο είναι το $1/4(-1+x)^2(1+x)(1+2x)^2$, δηλαδή η ιδιοτιμή 1 αλγεβρική πολλαπλότητα 2. Υπολογίζουμε τον ιδιόχωρο της ιδιοτιμής 1 ότι παράγεται από τα γραμμικά ανεξάρτητα διανύσματα $(0, 0, 0, 1, 1)^t, (1, 1, 1, 0, 0)^t$, συνεπώς δεν υπάρχει μοναδικό θετικό ιδιοδιάνυσμα της ιδιοτιμής 1. Ο πίνακας I_2 είναι στοχαστικός αλλά όχι θετικός.

Η λύση στα παραπάνω προβλήματα στον αλγόριθμο της Google είναι να αντικαταστήσουμε κάθε μηδενική γραμμή με $1/n$, έτσι ο πίνακας I_1 στο παράδειγμα IX.3.23 αντικαθίσταται από τον πίνακα

$$I'_n = I_1 = \begin{pmatrix} 0 & 1/2 & 1/3 \\ 1/2 & 0 & 1/3 \\ 1/2 & 1/2 & 1/3 \end{pmatrix}$$

Παρατηρήστε ότι σε ένα πίνακα που αναπαριστά το δίκτυο με τις δισεκατομμύρια σελίδες του το $1/n$ είναι πολύ μικρό. Ο πίνακας είναι τώρα στοχαστικός και το χαρακτηριστικό του πολυώνυμο

υπολογίζεται σε $1/12 + (7x)/12 + x^2/3 - x^3$ με ιδιοτιμές $1, -(1/2), -(1/6)$. Παρόλα αυτά ο πίνακας I_2 παραδείγματος IX.3.24 είναι στοχαστικός αλλά όχι θετικός. Ο πίνακας της Google είναι ο

$$M = (1 - p)A' + pB,$$

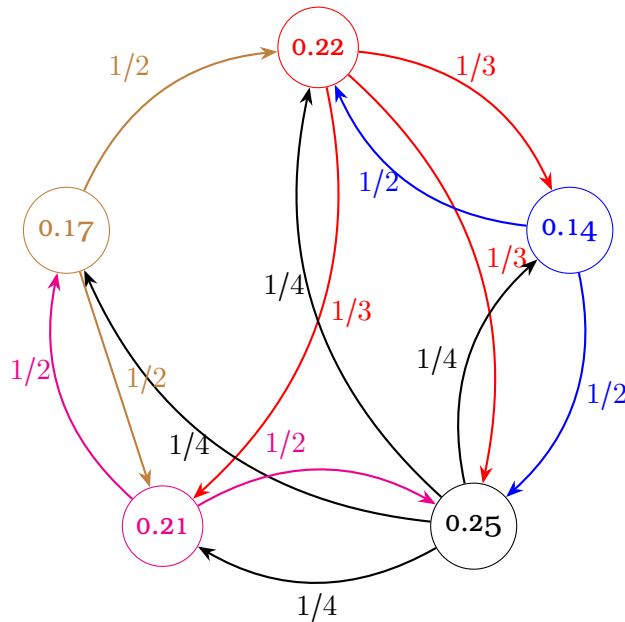
όπου ο πίνακας A' είναι ο πίνακας σημαντικότητας στον οποίο έχουμε αντικαταστήσει τις μη-δενικές στήλες με $1/n$, ενώ ο πίνακας $B = (b_{ij})$ είναι ο πίνακας $b_{ij} = 1/n$ και η σταθερά $p \in (0, 1)$ ονομάζεται συντελεστής απόσβεσης και μια τυπική τιμή του είναι $p = 0.15$.

Παρατήρηση IX.3.25. Ο πίνακας M είναι θετικός στοχαστικός οπότε σύμφωνα με το θεώρημα των Perron-Frobenius έχει μοναδικό θετικό διάνυσμα σταθερής κατάστασης. Ο υπολογισμός του είναι ένα δύσκολο υπολογιστικά πρόβλημα. Ο πίνακας του πραγματικού διαδικτύου έχει δισεκατομμύρια στήλες.

Παραδείγματα IX.3.26. Για τον πίνακα της IX.11 και $p = 0.15$ υπολογίζουμε ότι

$$M = (1 - p)A + pB = \begin{pmatrix} 0.03 & 0.455 & 0.2425 & 0.03 & 0.455 \\ 0.313333 & 0.03 & 0.2425 & 0.03 & 0.03 \\ 0.313333 & 0.455 & 0.03 & 0.455 & 0.03 \\ 0.313333 & 0.03 & 0.2425 & 0.03 & 0.455 \\ 0.03 & 0.03 & 0.2425 & 0.455 & 0.03 \end{pmatrix}$$

ο οποίος έχει διάνυσμα σταθερής κατάστασης το $s = (0.217624, 0.14383, 0.245504, 0.218156, 0.174886)^t$, δηλαδή η περισσότερο σημαντική σελίδα είναι η C. Ξαναγράφουμε το διάγραμμα με την βαθμολόγηση κάθε ιστοσελίδας στην εικόνα IX.3.



Σχήμα IX.3

Για τον πίνακα I_2 του παραδείγματος IX.3.24 υπολογίζουμε ότι

$$M = (1 - p)I_2 + pB = \begin{pmatrix} 0.03 & 0.455 & 0.455 & 0.03 & 0.03 \\ 0.455 & 0.03 & 0.455 & 0.03 & 0.03 \\ 0.455 & 0.455 & 0.03 & 0.03 & 0.03 \\ 0.03 & 0.03 & 0.03 & 0.03 & 0.88 \\ 0.03 & 0.03 & 0.03 & 0.88 & 0.03 \end{pmatrix}$$

και το σταθερό διάνυσμα είναι το $(1/5, 1/5, 1/5, 1/5, 1/5)^t$. Αυτό δεν είναι περίεργο, το διαδίκτυο που αντιστοιχεί στο γράφημα είναι συμμετρικό και καμία σελίδα δεν έχει καλύτερη βαθμολογία από τις άλλες.

Λύση της άσκησης I.1

1. Η Αθήνα δεν είναι η πρωτεύουσα της Ελλάδας
2. $3 + 5 \neq 8$
3. Υπάρχει $\epsilon > 0$ ώστε για κάθε $n_0(\epsilon)$ να υπάρχει $n > n_0(\epsilon)$ ώστε $-\epsilon \geq a_n$ ή $a_n \geq \epsilon$.
4. Τα μανταρίνια δεν είναι κίτρινα ή δεν έχουν κουκούτσια
5. Όλοι οι φοιτητές είναι πιο κοντοί από τα 2 μέτρα.

Λύση της άσκησης I.2

1. $\bar{p}$
2. $p \Rightarrow q$
3. $\bar{p} \Rightarrow \bar{q}$
4. $q \Rightarrow p$

Λύση της άσκησης I.3

1. Αν δεν μείνω σπίτι αύριο δεν θα βρέχει.
2. Αν δεν παίζουμε μπάλα αύριο τότε δεν θα έχει λιακάδα
3. Αν ένας θετικός ακέραιος έχει διαιρέτες διαφορετικούς από την μονάδα και τον εαυτό του, τότε δεν είναι πρώτος.

Λύση της άσκησης I.4

Πιάνουμε τον ένα στην τύχη και τον ρωτάμε: «αν ρώταγα τον άλλο ποιον δρόμο θα μου έδειχνε;» Ακολουθούμε τον άλλο δρόμο από αυτόν που μας έδειξε.

Λύση της άσκησης I.5

Θα δείξουμε ότι η παραπάνω σχέση είναι συμμετρική, ανακλαστική και μεταβατική.
Για να δείξουμε ότι είναι συμμετρική παρατηρούμε ότι αν

$$(a, b) \equiv (a', b') = 0 \Rightarrow ab' - ba' = 0 \Rightarrow a'b - b'a = 0 \Rightarrow (a', b') \equiv (a, b).$$

Για να δείξουμε ότι είναι ανακλαστική παρατηρούμε ότι

$$ab - ab = 0 \Rightarrow (a, b) \equiv (a, b).$$

Τέλος για να δείξουμε ότι είναι και μεταβατική παρατηρούμε ότι αν

$$(a_1, b_1) \equiv (a_2, b_2) \Rightarrow a_1 b_2 = b_1 a_2$$

και

$$(a_2, b_2) \equiv (a_3, b_3) \Rightarrow a_2 b_3 = b_2 a_3.$$

Άρα, πολλαπλασιάζοντας κατά μέλη έχουμε

$$a_1 b_2 a_2 b_3 = b_1 a_2 b_2 a_3 \Rightarrow a_1 b_3 = b_1 a_3 \Rightarrow (a_1, b_1) \equiv (a_3, b_3).$$

Λύση της άσκησης I.6

Η ζητούμενη πρόταση ισχύει για $k = 1$ αφού $1 = \frac{1(1+1)(2+1)}{6}$. Υποθέτουμε ότι ισχύει για $k = n$ και κάνοντας χρήση της υπόθεσης αυτής θα αποδείξουμε την αλήθεια της πρότασης για $k = n+1$. Έχουμε

$$1^2 + 2^2 + 3^2 + \dots + k^2 + (k+1)^2 = \frac{k(k+1)(2k+1)}{6} + (k+1)^2.$$

Στην παραπάνω ισότητα κάναμε χρήση της αλήθειας της πρότασης για $k = n$ προκειμένου να υπολογίσουμε το άθροισμα των k -πρώτων τετραγώνων. Τώρα το δεύτερο μέρος της ισότητας υπολογίζεται σε

$$\begin{aligned} \frac{k(k+1)(2k+1)}{6} + (k+1)^2 &= \frac{k(k+1)(2k+1) + 6(k+1)^2}{6} = \\ &= \frac{(k+1)(k+2)(2(k+1)+1)}{6}, \end{aligned}$$

και η αλήθεια της πρότασης για $n = k+1$ έχει αποδειχτεί.

Λύση της άσκησης I.7

Παρατηρούμε ότι η πρόταση είναι αληθής για $k = 1$ αφού

$$1 = 1 \cdot 1 = (1+1)! - 1.$$

Υποθέτουμε ότι η πρόταση είναι αληθής για $n = k$ και κάνοντας χρήση αυτού θα δείξουμε την αλήθεια της πρότασης για $n = k+1$. Πράγματι,

$$1 \cdot 1! + 2 \cdot 2! + 3 \cdot 3! + \dots + k \cdot k! + (k+1)! \cdot (k+1) = (k+1)! - 1 + (k+1)! \cdot (k+1).$$

Παρατηρήστε ότι στην παραπάνω πρόταση κάναμε χρήση της επαγωγικής υπόθεσης για να υπολογίσουμε το άθροισμα των k -πρώτων προσθετέων. Το δεύτερο μέρος της παραπάνω ισότητας είναι ίσο με

$$(k+1)! - 1 + (k+1)! \cdot (k+1) = (k+1)!(1+k+1) - 1 = (k+2)! - 1,$$

δηλαδή το ζητούμενο.

Λύση της άσκησης I.8

Παρατηρούμε ότι

$$\begin{aligned} \binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{(k-1)!(n-1-k)!} + \frac{(n-1)!}{k!(n-k)!} = \\ &= \frac{(n-1)!}{(k-1)!(n-1-k)!} \left(\frac{1}{k} + \frac{1}{n-k} \right) = \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k}, \end{aligned}$$

δηλαδή το ζητούμενο. Το δεύτερο κομμάτι της ασκήσης είναι άμεση εφαρμογή του πρώτου.

Λύση της άσκησης I.9

Παρατηρούμε ότι για $n = 1$ ο τύπος ισχύει:

$$(a+b)^n = \binom{n}{0}a + \binom{n}{n}b = a+b.$$

Στον παραπάνω τύπο παρατηρήστε ότι $\binom{n}{0} = \binom{n}{n} = 1$. Ας υποθέσουμε ότι ο τύπος ισχύει για $n = k$ και βάση αυτού θα αποδείξουμε ότι ο τύπος ισχύει για $n = k+1$. Έχουμε ότι

$$\begin{aligned} (a+b)^{k+1} &= (a+b)^k(a+b) = \\ &= \left(\sum_{m=0}^k \binom{k}{m} a^m b^{k-m} \right) (a+b) = \\ &= a^{k+1} + \binom{k}{1} a^k b + \binom{k}{2} a^{k-1} b^2 + \dots + \binom{k}{k-2} a^3 b^{k-2} + \binom{k}{k-1} a^2 b^{k-1} + \binom{k}{k} a b^k + \\ &\quad + a^k b + \binom{k}{1} a^{k-1} b^2 + \dots + \binom{k}{k-2} a^2 b^{k-1} + \binom{k}{k-1} a b^k + \binom{k}{k} b^{k+1} = \\ &= a^{k+1} + \sum_{m=1}^k \left(\binom{k}{m} + \binom{k}{m-1} \right) a^{k+1-m} b^m + b^{k+1} = \\ &= \sum_{m=0}^{k+1} \binom{k+1}{m} a^{k+1-m} b^m. \end{aligned} \tag{I.5}$$

Λύση της άσκησης I.10

Παρατηρούμε ότι η ζητούμενη ισότητα ισχύει για $k = 1$ αφού

$$1^2 = (-1)^0 \frac{1(1+1)}{2}.$$

Υποθέτουμε ότι είναι αληθής για $n = k$ και κάνοντας χρήση αυτού θα δείξουμε την αλήθεια της πρότασης για $n = k+1$. Παρατηρούμε ότι

$$1^2 - 2^2 + 3^2 - 4^2 + \dots + (-1)^{k-1} k^2 + (-1)^k (k+1)^2 = (-1)^{k-1} \frac{k(k+1)}{2} + (-1)^k (k+1)^2.$$

Στην παραπάνω ισότητα κάναμε χρήση της επαγωγικής υπόθεσης για να υπολογίσουμε το άθροισμα των k -πρώτων προσθετέων. Παρατηρούμε ότι το δεύτερο μέρος της ισότητας είναι ίσο με

$$\begin{aligned} (-1)^{k-1} \frac{k(k+1)}{2} + (-1)^k (k+1)^2 &= (-1)^k \left(-\frac{k^2+k}{2} + \frac{2k^2+4k+2}{2} \right) = \\ &= (-1)^k \frac{k^2+3k+2}{2} = (-1)^k \frac{(k+1)(k+2)}{2}, \end{aligned}$$

δηλαδή το ζητούμενο.

Λύση της άσκησης I.11

Για $n = 1$ έχουμε ότι $f(1) \geq 1$, αφού το $f(1) \in \mathbb{N}$. Ας υποθέσουμε ότι για $n = k$ έχουμε $f(k) \geq k$. Έχουμε ότι $f(k+1) > f(k) \geq k$, άρα $f(k+1) \geq k$.

Λύση της άσκησης I.12

1. Σωστό, αφού αν $x = \frac{a}{b}$ και $y = \frac{c}{d}$ είναι δύο ρητοί αριθμοί ($a, b, c, d \in \mathbb{Z}$) τότε και το άθροισμα τους θα είναι ρητός $x + y = \frac{ad+cb}{bd}$.
2. Σωστό. Υποθέτουμε ότι $x \in \mathbb{Q}$ και $y \in \mathbb{R} \setminus \mathbb{Q}$ και ας υποθέσουμε, για να καταλήξουμε σε άτοπο, ότι $x + y \in \mathbb{Q}$. Τότε $y = (x + y) + (-x)$ είναι το άθροισμα δύο ρητών και θα έπρεπε το y να είναι ρητός, άτοπο.
3. Λάθος, αφού για παράδειγμα ο $\sqrt{2}$ είναι άρρητος, άλλα και ο $1 - \sqrt{2}$ είναι άρρητος (γιατί?). Το άθροισμα τους είναι όμως $1 \in \mathbb{Q}$.
4. Λάθος, αφού για παράδειγμα ο $\sqrt{2}$ είναι άρρητος, και έχουμε ότι $\sqrt{2} + \sqrt{2} = 2\sqrt{2}$ ο οποίος είναι άρρητος.
5. Σωστό, αφού αν $x \in \mathbb{R}$ μπορούμε να θεωρήσουμε την ακολουθία των ρητών προσεγγύσεων του $a_n = \frac{[10^n x]}{10^n}$. Παρατηρούμε ότι $a_n \in \mathbb{Q}$ και ότι

$$x - a_n < \frac{1}{10^n} \rightarrow 0.$$

6. Σωστό, έστω $x \in \mathbb{R}$, θεωρούμε το $x - \sqrt{2}$ και διαλέγουμε μία ακολουθία ρητών (a_n) με $a_n \rightarrow x - \sqrt{2}$. Η ακολουθία (b_n) με γενικό όρο $b_n = a_n + \sqrt{2}$ είναι μία ακολουθία αρρήτων η οποία συγκλίνει στο x .
7. Έστω $x, y \in \mathbb{R}$ με $x < y$. Υπάρχει ακολουθία άρρητων αριθμών (a_n) η οποία να συγκλίνει στο $(x + y)/2$ ισχύει:

$$x < \frac{x+y}{2} < y$$

Από τον ορισμό σύγκλισης της ακολουθίας $a_n \rightarrow \frac{x+y}{2}$ έχουμε ότι για $\epsilon = \frac{x+y}{4}$, όλοι οι όροι της ακολουθίας από ένα n_0 και μετά θα είναι μέσα στο διάστημα (x, y) , από όπου προκύπτει το ζητούμενο.

Λύση της άσκησης I.13

Αν a, b είναι ρητοί αριθμοί τότε ο $(a + b)/2$ είναι ρητός και επιπλέον

$$a \leq \frac{a+b}{2} \leq b.$$

Λύση της άσκησης I.14

Η αρχή του Αρχιμήδη I.6.16 εξασφαλίζει ότι υπάρχει φυσικός αριθμός ώστε $n(b - a) \geq 1$. Άρα αφού οι αριθμοί na, nb απέχουν περισσότερο από 1 υπάρχει ανάμεσα τους ένας ακέραιος k . Τέλος έχουμε

$$a \leq \frac{k}{n} \leq b,$$

δηλαδή το ζητούμενο.

Λύση της άσκησης I.15

Θεωρούμε δύο αρχικά ανοιχτά τμήματα A_1, A_2 τα οποία είναι μή κενά και είναι θετικοί αριθμοί δηλαδή

$$\{x \in \mathbb{Q} : x < 0\} \not\subset A_i$$

Δηλαδή υπάρχουν $a_1 \in A_1$ και $a_2 \in A_2$ με $a_i > 0$. Άρα στο γινόμενο $A_1 A_2$ υπάρχει το θετικό στοιχείο $a_1 a_2 > 0$. Επιπλέον $\{x \in \mathbb{Q} : x < 0\} \subset A_1 A_2$ από όπου προκύπτει το ζητούμενο.

Λύση της άσκησης I.16

Έχουμε

$$\frac{1}{a+ib} = \frac{a-ib}{(a+ib)(a-ib)} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} - i \frac{b}{a^2+b^2}.$$

Λύση της άσκησης I.17

Θέτουμε $\Delta = 1^2 - 12 = -11$, οπότε $\sqrt{-11} = \sqrt{11}i$. Οι ρίζες δίνονται από τους τύπους

$$\rho_1 = -\frac{1}{2} + \frac{\sqrt{11}}{2}i, \quad \rho_2 = -\frac{1}{2} - \frac{\sqrt{11}}{2}i.$$

Λύση της άσκησης I.18

Έχουμε ότι $|1+i| = \sqrt{2}$. Άρα

$$1+i = \sqrt{2} \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i \right) = \sqrt{2} \left(\cos\left(\frac{\pi}{4}\right) + i \sin\left(\frac{\pi}{4}\right) \right).$$

Συνεπώς οι 10-ες ρίζες του $1+i$ δίνονται από τον τύπο

$$z_k = \sqrt[10]{2} \left(\cos\left(\frac{2\pi k + \frac{\pi}{4}}{10}\right) + i \sin\left(\frac{2\pi k + \frac{\pi}{4}}{10}\right) \right),$$

όπου $k = 0, \dots, 9$.

Λύση της άσκησης I.19

Αν $x, y \in \mu(n)$ τότε $x^n = y^n = 1 \Rightarrow (xy)^n = x^n y^n = 1$. Η μονάδα 1 είναι σαφώς n -ρίζα της μονάδας, δηλαδή $1 \in \mu(n)$. Τέλος αν $x^n = 1$ τότε και $(1/x)^n = 1/x^n = 1$.

Λύση της άσκησης I.20

Η z^n απεικονίζει το D εντός του D , αφού αν $|z| < 1$ τότε και $|z^n| = |z|^n < 1$. Αν $a \in D$, τότε η εξίσωση $x^n = a$ έχει n -το πλήθος ρίζες στο $\mathbb{C}$ και όλες είναι εντός του D , αφού αν $|z| \geq 1$ τότε και $|a| = |z^n| = |z|^n \geq 1$. Δηλαδή η z^n είναι επί του D . Αφού για κάθε $a \neq 0$ υπάρχουν n το πλήθος στοιχεία z_i εντός του D ώστε $z_i^n = a$ δεν είναι δυνατόν να είναι η συνάρτηση ένα προς ένα.

Γεωμετρικά η z^n χωρίζει τον δίσκο D σε n φέτες της μορφής $D_j = \{z \in D, 2\pi(j-1) \leq \text{Arg}(z) < 2\pi j\}$ και η συνάρτηση $z^n|_{D_j}$ είναι ένα προς ένα και επί, δηλαδή η z^n απλώνει την φέτα D_j πάνω σε ολόκληρο το δίσκο.

Λύση της άσκησης I.21

Οι n -ρίζες της μονάδας είναι οι ζ_1^k , όπου $\zeta_1 = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right)$ και $k = 0, \dots, n-1$. Έτσι

$$1 + \zeta_1 + \zeta_1^2 + \dots + \zeta_1^{n-1} = \frac{\zeta_1^n - 1}{\zeta_1 - 1} = 0.$$

Λύση της άσκησης I.22

Υπενθυμίζουμε ότι $z \in \mathbb{C}$ είναι πραγματικός αν και μόνο $\bar{z} = z$. Θεωρούμε το πολυώνυμο $f(x) = \sum_{v=0}^n a_v x^v \in \mathbb{R}[x]$. Αν $\rho \in \mathbb{C}$ είναι μια ρίζα του $f(x)$ δηλαδή $f(\rho) = 0$, τότε

$$0 = f(\rho) = \overline{f(\rho)} = \sum_{v=0}^n \overline{a_v \rho^v} = \sum_{v=0}^n \overline{a_v} \overline{\rho^v} = \sum_{v=0}^n \overline{a_v} \overline{\rho}^v = f(\bar{\rho}),$$

δηλαδή και η $f(\bar{\rho}) = 0$.

Λύση της άσκησης I.23

Παρατηρούμε ότι

$$\cos^2(t) + \sin^2(t) = |\cos(t) + i \sin(t)|^2 = e^{it} \overline{e^{it}} = e^{it} e^{-it} = e^{it} e^{-it} = 1$$

Λύση της άσκησης I.24

Υπολογίζουμε ότι

$$\sum_{v=0}^n e^{i v a} = \frac{e^{i(n+1)a} - 1}{e^{ia} - 1} = \frac{(e^{i(n+1)a} - 1)(e^{-ia} - 1)}{(e^{ia} - 1)(e^{-ia} - 1)}.$$

Το ζητούμενο άθροισμα είναι το πραγματικό μέρος του αθροίσματος της παραπάνω εξίσωσης. Υπολογίζουμε τον παρονομαστή ως

$$(e^{ia} - 1)(e^{-ia} - 1) = e^{ia/2} (e^{ia/2} - e^{-ia/2}) e^{-ia/2} (e^{-ia/2} - e^{ia/2}) = -4 \sin^2(a/2)$$

και τον αριθμητή ως

$$\begin{aligned} (e^{i(n+1)a} - 1)(e^{-ia} - 1) &= e^{i(n+1)a/2} (e^{i(n+1)a/2} - e^{-i(n+1)a/2}) e^{-ia/2} (e^{-ia/2} - e^{ia/2}) \\ &= e^{ina/2} (-2i) \sin((n+1)a/2) (-2i) \sin(a/2) \\ &= -4(\cos(na/2) + i \sin(na/2)) \sin((n+1)a/2) \sin(a/2) \end{aligned}$$

από όπου προκύπτει το ζητούμενο.

Λύση της άσκησης II.1

Έστω ο πίνακας $A = (a_{ij})$ τον οποίο γράφουμε ως $A = (A + A^t)/2 + (A - A^t)/2$. Έχουμε ότι $(A + A^t)^t = A^t + (A^t)^t = A^t + A$ ενώ $(A - A^t)^t = A^t - A = -(A - A^t)$.

Λύση της άσκησης II.2

Θα το αποδείξουμε με επαγωγή. Για $r = 2$ το ζητούμενο αποτέλεσμα είναι απλά ο πολλαπλασιασμός δύο πινάκων. Υποθέτουμε την αλήθεια του ζητουμένου για το γινόμενο $r - 1$ πινάκων και έχουμε ότι αν $(b_{ij}) = A^{(1)} \dots A^{(r-1)}$, τότε υπολογίζουμε το γινόμενο $(b_{ij}) A^{(r)}$

$$a_{ij} = \sum_{v_r=1}^{\ell_r} b_{i v_r} a_{v_r, j}^{(r)} = \sum_{v_1=1}^{\ell_1} \sum_{v_2=1}^{\ell_2} \dots \sum_{v_r=1}^{\ell_r} a_{i, v_1}^{(1)} a_{v_1, v_2}^{(2)} \dots a_{v_{r-1}, j}^{(r)}.$$

Λύση της άσκησης II.3

Γράφουμε $A = (a_{ij})$, $b = (b_{ij})$ και έχουμε ότι ο πίνακας $AB = (c_{ij})$ δίνεται από τον τύπο

$$c_{ij} = \sum_{v=1}^m a_{iv} b_{vj}.$$

Οι ανάστροφοι πίνακες δίνονται από τα στοιχεία $A^t = (a'_{ij})$, $B^t = (b'_{ij})$, $(AB)^t = (c'_{ij})$ όπου $a'_{ij} = a_{ji}$, $b'_{ij} = b_{ji}$, $c'_{ij} = c_{ji}$ και η παραπάνω σχέση πολλαπλασιασμού δίνει

$$c'_{ij} = c_{ji} = \sum_{v=1}^m a_{jv} b_{vi} = \sum_{v=1}^m a'_{v,j} b'_{i,v} = \sum_{v=1}^m b'_{i,v} a'_{v,j},$$

δηλαδή το ζητούμενο.

Λύση της άσκησης II.4

Ο πίνακας A είναι αντιστρέψιμος, δηλαδή υπάρχει πίνακας A^{-1} ώστε

$$AA^{-1} = A^{-1}A = \mathbb{I}_n.$$

Εφαρμόζουμε τον ανάστροφο στην παραπάνω σχέση για να πάρουμε:

$$(A^{-1})^t A^t = A^t (A^{-1})^t = \mathbb{I}_n.$$

Ο πίνακας $(A^{-1})^t$ είναι ένας αντίστροφος και από την μοναδικότητα του αντιστρόφου έχουμε $(A^t)^{-1} = (A^{-1})^t$.

Λύση της άσκησης II.5

Υπάρχουν αρκετοί τρόποι να αντιμετωπιστεί το πρόβλημα αυτό. Είναι εφικτό να θεωρήσουμε τον πίνακα

$$B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

και να λύσουμε το σύστημα

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ισοδύναμα

$$\begin{array}{rcl} ax & +bz & = 1 \\ & ay & +bw = 0 \\ cx & +dz & = 0 \\ & cy & +dw = 1 \end{array}$$

Ας υποθέσουμε ότι $a \neq 0$ και ας πολλαπλασιάσουμε την πρώτη και δεύτερη εξίσωση και ας τις προσθέσουμε στην τρίτη και τέταρτη εξίσωση για να πάρουμε

$$\begin{array}{rcl} ax & +bz & = 1 \\ & ay & +bw = 0 \\ \left(\frac{ad-bc}{a}\right)z & & = -\frac{c}{a} \\ & \left(\frac{ad-bc}{a}\right)w & = 1 \end{array}$$

Αν κάνουμε και την υπόθεση ότι $D = ad - bc \neq 0$ τότε καταλήγουμε στο ότι $z = -c/D$ και $w = a/D$. Αντικαθιστώντας την τιμή του z στην τρίτη εξίσωση και την τιμή του w στην δεύτερη έχουμε ότι $x = d/D$ και $y = -b/D$.

Αν το $a = 0$ τότε η πρώτη εξίσωση δίνει ότι $z = 1/b = -c/D$ στην περίπτωση που $b \neq 0$ ενώ αν $b = 0$ η εξίσωση είναι αδύνατη. Ομοίως η δεύτερη εξίσωση δίνει $w = 0 = a/D$. Αντικαθιστώντας τις τιμές

αυτές στην τρίτη και τέταρτη εξίσωση έχουμε ότι αν $c = 0$ η τέταρτη εξίσωση είναι αδύνατη ενώ αν $c \neq 0$ τότε $y = 1/c = -b/D$ και επίσης $z = -c/D = 1/b$. Δηλαδή ακόμα και αν $a = 0$ ο αντίστροφος πίνακας υπολογίζεται ακριβώς με τους ίδιους τύπους όπως και στην περίπτωση $a \neq 0$.

Ένας δεύτερος τρόπος θα ήταν να παρατηρήσουμε (κάνοντας τις πράξεις) ότι

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Αν λοιπόν $D = ad - bc = 0$ ο πίνακας A αποκλείεται να είναι αντιστρέψιμος, ενώ αν $D = ad - bc \neq 0$ τότε ο πίνακας

$$A' = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

είναι έχει την ζητούμενη ιδιότητα πολλαπλασιαζόμενος από τα αριστερά με τον A .

Και στις δύο μεθόδους μπορούμε να επαληθεύσουμε ότι ο πίνακας A' ικανοποιεί και την σχέση $A'A = \mathbb{I}_2$. Αυτό όπως θα αποδείξουμε είναι ένα γενικότερο φαινόμενο δηλαδή αν $AA' = \mathbb{I}_n$ τότε και $A'A = \mathbb{I}_n$.

Λύση της άσκησης II.6

Για το πρώτο ερώτημα παρατηρούμε ότι αν $A = (a_{ij})$, $B = (b_{ij})$ τότε

$$\text{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji} = \sum_{j=1}^n \sum_{i=1}^n b_{ji} a_{ij} = \text{tr}(BA).$$

Για το δεύτερο ερώτημα έχουμε

$$\text{tr}(QAQ^{-1}) = \text{tr}(AQ^{-1}Q) = \text{tr}(A).$$

Λύση της άσκησης II.7

Θεωρούμε τους πίνακες της μορφής $E_{ij} = (b_{k,l}^{(i,j)})$ οι οποίοι έχουν μονάδα στην i -γραμμή και στην j -στήλη και παντού αλλού 0 δηλαδή $b_{k,l}^{(i,j)} = \delta_{i,k} \delta_{j,l}$. Παρατηρούμε ότι $E_{i_1,j_1} E_{i_2,j_2} = \delta_{j_1,i_2} E_{i_1,j_2}$. Πράγματι,

$$\sum_{v=1}^n b_{k,v}^{(i_1,j_1)} b_{v,l}^{(i_2,j_2)} = \sum_{v=1}^n \delta_{k,i_1} \delta_{v,j_1} \delta_{v,i_2} \delta_{l,j_2} = \delta_{k,i_1} \delta_{j_1,i_2} \delta_{l,j_2} = \delta_{j_1,i_2} E_{i_1,j_2}.$$

Ο πίνακας $A = (a_{ij})$ γράφεται ως

$$A = \sum_{k,l=1}^n a_{kl} E_{kl}$$

και αν αντιμετωπίζεται με κάθε πίνακα έχουμε

$$AE_{ij} = E_{ij}A \text{ ισοδύναμα } \sum_{k,l=1}^n a_{kl} E_{kl} E_{ij} = E_{ij} \sum_{k,l=1}^n a_{kl} E_{kl}$$

από όπου έχουμε ότι

$$\sum_{k,l=1}^n a_{kl} \delta_{li} E_{kj} = \sum_{k,l=1}^n a_{kl} \delta_{jk} E_{il} \Leftrightarrow \sum_{k=1}^n a_{ki} E_{kj} = \sum_{l=1}^n a_{jl} E_{il}$$

από όπου έχουμε ότι $a_{ii} = a_{jj}$ για όλα τα $1 \leq i \leq j \leq n$ και $a_{ki} = 0$ όταν $k \neq i$. Το αντίστροφο ότι οι πίνακες $\lambda \mathbb{I}_n$ αντιμετωπίζονται με κάθε πίνακα είναι προφανές.

Λύση της άσκησης II.8

Αν έχουμε ότι $\lambda_1 \lambda_2 \cdots \lambda_n = 0$ τότε τουλάχιστον ένα από τα $\lambda_1, \dots, \lambda_n$ είναι μηδενικό. Έστω $\lambda_{i_0} = 0$. Τότε για τον διαγώνιο πίνακα $D = (\mu_1, \dots, \mu_n)$ με $\mu_i = \delta_{i, i_0}$ έχουμε

$$\text{diag}(\lambda_1, \dots, \lambda_n) D = \mathbf{0},$$

και αν ο πίνακας $\text{diag}(\lambda_1, \dots, \lambda_n)$ ήταν αντιστρέψιμος και πολλαπλασιάζαμε από αριστερά με τον αντίστροφό του θα είχαμε ότι και ο D θα ήταν μηδενικός, άτοπο.

Λύση της άσκησης II.9

Έστω $A = (a_{ij}), B = (b_{ij})$. Το γινόμενο έχει στοιχεία $AB = (c_{ij})$. Έχουμε ότι

$$c_{ij} = \sum_{v=1}^p a_{iv} b_{vj} = \sum_{v=1}^{p_1} a_{iv} b_{vj} + \sum_{v=p_1+1}^{p_1+p_2} a_{iv} b_{vj}.$$

Ας υποθέσουμε ότι $i \leq n_1, j \leq p_1$ τότε το παραπάνω άθροισμα γράφεται ως

$$c_{ij} = \sum_{v=1}^{p_1} a_{iv}^{(11)} b_{vj}^{(12)} + \sum_{v=1}^{p_2} a_{iv}^{(12)} b_{vj}^{(21)}$$

Δηλαδή το ζητούμενο. Με ανάλογο τρόπο αποδεικνύονται οι άλλες τρεις περιπτώσεις για τα i, j . Τέλος υπολογίζουμε ότι

$$J^2 = -\mathbb{I}_{2n}.$$

Λύση της άσκησης III.1

Κατασκευάζουμε τον επαυξημένο πίνακα:

$$\begin{aligned} & \left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 5 \\ 2 & -1 & 3 & -2 & 8 \\ 3 & 1 & 1 & 4 & 4 \\ 4 & -2 & 2 & -1 & 2 \end{array} \right) \begin{array}{l} \leftarrow_{+}^{-2} \\ \leftarrow_{+}^{-3} \\ \leftarrow_{+}^{-4} \end{array} \sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 5 \\ 0 & -5 & 5 & -8 & -2 \\ 0 & -5 & 4 & -5 & -11 \\ 0 & -10 & 6 & -13 & -18 \end{array} \right) \begin{array}{l} \leftarrow_{+}^{-1} \\ \leftarrow_{+}^{-2} \end{array} \sim \\ & \left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 5 \\ 0 & -5 & 5 & -8 & -2 \\ 0 & 0 & -1 & 3 & -9 \\ 0 & 0 & -4 & 3 & -14 \end{array} \right) \begin{array}{l} \leftarrow_{+}^{-4} \\ \leftarrow_{+} \end{array} \sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 5 \\ 0 & -5 & 5 & -8 & -2 \\ 0 & 0 & -1 & 3 & -9 \\ 0 & 0 & 0 & -9 & 22 \end{array} \right) \begin{array}{l} \cdot -1/5 \\ \cdot -1 \\ \cdot -1/9 \end{array} \sim \\ & \left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 5 \\ 0 & 1 & -1 & 8/5 & 2/5 \\ 0 & 0 & 1 & -3 & 9 \\ 0 & 0 & 0 & 1 & -22/9 \end{array} \right) \begin{array}{l} \leftarrow_{+} \\ \leftarrow_{+} \\ \leftarrow_{+} \end{array} \sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 37/3 \\ 0 & 1 & -1 & 0 & 194/45 \\ 0 & 0 & 1 & 0 & 5/3 \\ 0 & 0 & 0 & 1 & -22/9 \end{array} \right) \begin{array}{l} \leftarrow_{+} \\ \leftarrow_{+} \\ \leftarrow_{+} \end{array} \sim \\ & \left(\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 14 \\ 0 & 1 & 0 & 0 & 269/45 \\ 0 & 0 & 1 & 0 & 5/3 \\ 0 & 0 & 0 & 1 & -22/9 \end{array} \right) \begin{array}{l} \leftarrow_{+} \\ \leftarrow_{-2} \end{array} \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 92/45 \\ 0 & 1 & 0 & 0 & 269/45 \\ 0 & 0 & 1 & 0 & 5/3 \\ 0 & 0 & 0 & 1 & -22/9 \end{array} \right) \end{aligned}$$

οπότε έχουμε την μοναδική λύση $(x, y, z, w) = (92/45, 269/45, 5/3, -22/9)$.

Λύση της άσκησης III.2

Το παραπάνω σύστημα έχει επαυξημένο πίνακα στο οποίο εκτελούμε την μέθοδο της απαλοιφής

$$\begin{pmatrix} a & 1 & 1 & | & 1 \\ a & a & 1 & | & 1 \\ 1 & 1 & a & | & 1 \end{pmatrix} \xleftarrow{\quad} \sim \begin{pmatrix} 1 & 1 & a & | & 1 \\ a & a & 1 & | & 1 \\ a & 1 & 1 & | & 1 \end{pmatrix} \xleftarrow{\begin{matrix} \boxed{-a} \\ \boxed{+} \\ \boxed{+} \end{matrix}} \sim$$

$$\begin{pmatrix} 1 & 1 & a & | & 1 \\ 0 & 0 & -a^2+1 & | & -a+1 \\ 0 & -a+1 & -a^2+1 & | & -a+1 \end{pmatrix} \xleftarrow{\begin{matrix} \boxed{-1} \\ \boxed{+} \end{matrix}} \sim \begin{pmatrix} 1 & 1 & a & | & 1 \\ 0 & 0 & -a^2+1 & | & -a+1 \\ 0 & -a+1 & 0 & | & 0 \end{pmatrix}$$

Διακρίνουμε τώρα τις περιπτώσεις:

Αν $a = 1$ τότε ο τελευταίος πίνακας γράφεται:

$$\begin{pmatrix} 1 & 1 & 1 & | & 1 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$

και οδηγεί στις παρακάτω λύσεις του συστήματος

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad x, y \in \mathbb{F}.$$

Αν το $a = -1$ τότε ο πίνακας γράφεται στην μορφή

$$\begin{pmatrix} 1 & 1 & -1 & | & 1 \\ 0 & 0 & 0 & | & 2 \\ 0 & 2 & 0 & | & 0 \end{pmatrix}$$

το οποίο είναι αδύνατο. Και αν $a \neq \pm 1$ τότε συνεχίζουμε την απαλοιφή:

$$\begin{pmatrix} 1 & 1 & a & | & 1 \\ 0 & 0 & -a^2+1 & | & -a+1 \\ 0 & -a+1 & 0 & | & 0 \end{pmatrix} \begin{matrix} | \cdot \frac{1}{1-a^2} \\ | \cdot \frac{1}{1-a} \end{matrix} \sim \begin{pmatrix} 1 & 1 & a & | & 1 \\ 0 & 0 & 1 & | & \frac{1}{1+a} \\ 0 & 1 & 0 & | & 0 \end{pmatrix} \xleftarrow{\quad}$$

$$\begin{pmatrix} 1 & 1 & a & | & 1 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & \frac{1}{1+a} \end{pmatrix} \xleftarrow{\begin{matrix} \boxed{+} \\ \boxed{-a} \end{matrix}} \xleftarrow{\begin{matrix} \boxed{-1} \\ \boxed{+} \end{matrix}} \sim \begin{pmatrix} 1 & 0 & 0 & | & \frac{2+a}{a} \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & \frac{1}{1+a} \end{pmatrix}$$

από όπου προκύπτει μοναδική λύση του συστήματος $x = \frac{2+a}{s}, y = 0, z = \frac{1}{1+a}$.

Λύση της άσκησης III.3

Το παραπάνω σύστημα αντιστοιχεί στον παρακάτω επαυξημένο πίνακα στον οποίο κάνουμε απαλοιφή:

$$\begin{pmatrix} a & 1 & 1 & | & 1 \\ 1 & a & 1 & | & a \\ 1 & 1 & a & | & a^2 \end{pmatrix} \xleftarrow{\quad} \sim \begin{pmatrix} 1 & 1 & a & | & a^2 \\ 1 & a & 1 & | & a \\ a & 1 & 1 & | & 1 \end{pmatrix} \xleftarrow{\begin{matrix} \boxed{-1} \\ \boxed{+} \\ \boxed{+} \end{matrix}} \sim$$

$$\begin{pmatrix} 1 & 1 & a & | & a^2 \\ 0 & a-1 & -a+1 & | & -a^2+a \\ 0 & -a+1 & -a^2+1 & | & -a^3+1 \end{pmatrix} \xleftarrow{\begin{matrix} \boxed{-1} \\ \boxed{+} \end{matrix}} \sim \begin{pmatrix} 1 & 1 & a & | & a^2 \\ 0 & a-1 & -a+1 & | & -a^2+a \\ 0 & 0 & -a^2-a+2 & | & -a^3-a^2+a+1 \end{pmatrix}$$

Έχουμε ότι $-a^2 - a + 2 = (1-a)(a+2)$. Και πάλι διακρίνουμε περιπτώσεις:

Αν $a = 1$ τότε το σύστημα γράφεται ως

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

και οδηγεί στις παρακάτω λύσεις του συστήματος

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad x, y \in \mathbb{F}.$$

Αν $a = -2$ τότε έχουμε τον πίνακα

$$\left(\begin{array}{ccc|c} 1 & 1 & -2 & 4 \\ 0 & -3 & 1 & -6 \\ 0 & 0 & 0 & 3 \end{array} \right)$$

ο οποίος οδηγεί σε αδύνατο σύστημα.

Τέλος για $a \neq 1, -2$ συνεχίζουμε με την απαλοιφή

$$\left(\begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & a-1 & -a+1 & -a^2+a \\ 0 & 0 & -a^2-a+2 & -a^3-a^2+a+1 \end{array} \right) \begin{array}{l} \\ \cdot \frac{1}{a-1} \\ \cdot \frac{1}{(1-a)(a+2)} \end{array} \sim \left(\begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & 1 & -1 & -a \\ 0 & 0 & 1 & \frac{(1+a)^2}{a+2} \end{array} \right) \begin{array}{l} \leftarrow + \\ \\ \leftarrow -1 \end{array}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & a+1 & a^2+a \\ 0 & 1 & -1 & -a \\ 0 & 0 & 1 & \frac{(1+a)^2}{a+2} \end{array} \right) \begin{array}{l} \leftarrow + \\ \leftarrow + \\ \leftarrow - \end{array} \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{-a-1}{a+2} \\ 0 & 1 & 0 & \frac{1}{a+2} \\ 0 & 0 & 1 & \frac{(1+a)^2}{a+2} \end{array} \right)$$

από όπου προκύπτει μοναδική λύση $x = \frac{-a-1}{a+2}$, $y = \frac{1}{a+2}$, $z = \frac{(1+a)^2}{a+2}$.

Λύση της άσκησης III.4

θεωρούμε τον πίνακα

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 7 & 1 & 0 & 0 & 0 \\ 5 & 4 & 1 & 10 & 0 & 1 & 0 & 0 \\ 9 & -1 & 0 & 3 & 0 & 0 & 1 & 0 \\ 10 & 0 & 13 & & 1 & 0 & 0 & 1 \end{array} \right)$$

Πολλαπλασιάζουμε την πρώτη γραμμή με -5 και την προσθέτουμε στην δεύτερη την πρώτη γραμμή με -9 και την προσθέτουμε στην τρίτη, την πρώτη γραμμή με -10 και την προσθέτουμε στην τέταρτη για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 7 & 1 & 0 & 0 & 0 \\ 0 & -11 & -24 & -25 & -5 & 1 & 0 & 0 \\ 0 & -28 & -45 & -60 & -9 & 0 & 1 & 0 \\ 0 & -30 & -37 & -69 & 10 & 0 & 0 & 1 \end{array} \right)$$

Στην συνέχεια πολλαπλασιάζουμε την δεύτερη γραμμή με $-\frac{28}{11}$ και προσθέτουμε στην τρίτη, την δεύτερη γραμμή με $-\frac{30}{11}$ και προσθέτουμε στην τέταρτη για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 7 & 1 & 0 & 0 & 0 \\ 0 & -11 & -24 & -25 & -5 & 1 & 0 & 0 \\ 0 & 0 & \frac{177}{11} & \frac{40}{11} & \frac{41}{11} & -\frac{28}{11} & 1 & 0 \\ 0 & 0 & \frac{313}{11} & -\frac{9}{11} & \frac{40}{11} & -\frac{30}{11} & 0 & 1 \end{array} \right)$$

Πολλαπλασιάζουμε την τρίτη γραμμή με $-\frac{311}{177}$ για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 7 & 1 & 0 & 0 & 0 \\ 0 & -11 & -24 & -25 & -5 & 1 & 0 & 0 \\ 0 & 0 & \frac{177}{11} & \frac{40}{11} & \frac{41}{11} & -\frac{28}{11} & 1 & 0 \\ 0 & 0 & 0 & -\frac{1283}{177} & -\frac{523}{177} & \frac{314}{177} & -\frac{313}{177} & 1 \end{array} \right)$$

Πολλαπλασιάζουμε την τελευταία γραμμή με $-\frac{177}{1283}$ για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 7 & 1 & 0 & 0 & 0 \\ 0 & -11 & -24 & -25 & -5 & 1 & 0 & 0 \\ 0 & 0 & \frac{177}{11} & \frac{40}{11} & \frac{41}{11} & -\frac{28}{11} & 1 & 0 \\ 0 & 0 & 0 & 1 & \frac{11}{1283} & -\frac{314}{1283} & \frac{313}{1283} & -\frac{177}{1283} \end{array} \right)$$

Στην συνέχεια πολλαπλασιάζουμε την τελευταία γραμμή με $-\frac{40}{11}$ και την προσθέτουμε στην τρίτη, την τελευταία γραμμή με 25 και την προσθέτουμε στην δεύτερη και την τελευταία με -7 και την προσθέτουμε στην πρώτη για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 0 & -\frac{2378}{1283} & \frac{2198}{1283} & -\frac{2191}{1283} & \frac{1239}{1283} \\ 0 & -11 & -24 & 0 & \frac{6660}{1283} & -\frac{6567}{1283} & \frac{7825}{1283} & -\frac{4425}{1283} \\ 0 & 0 & \frac{177}{11} & 0 & \frac{31683}{1283} & -\frac{2124}{1283} & \frac{1593}{1283} & \frac{7080}{1283} \\ 0 & 0 & 0 & 1 & \frac{14113}{1283} & -\frac{1283}{1283} & \frac{14113}{1283} & \frac{14113}{1283} \end{array} \right)$$

Πολλαπλασιάζουμε την τρίτη γραμμή με $\frac{11}{177}$ για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 5 & 0 & -\frac{2378}{1283} & \frac{2198}{1283} & -\frac{2191}{1283} & \frac{1239}{1283} \\ 0 & -11 & -24 & 0 & \frac{6660}{1283} & -\frac{6567}{1283} & \frac{7825}{1283} & -\frac{4425}{1283} \\ 0 & 0 & 1 & 0 & \frac{179}{1283} & -\frac{132}{1283} & \frac{9}{1283} & \frac{40}{1283} \\ 0 & 0 & 0 & 1 & \frac{523}{1283} & -\frac{314}{1283} & \frac{313}{1283} & -\frac{177}{1283} \end{array} \right)$$

Τώρα πολλαπλασιάζουμε την τρίτη γραμμή με 24 και την προσθέτουμε στην δεύτερη και την τρίτη γραμμή με -5 και την προσθέτουμε στην πρώτη για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 0 & 0 & -\frac{3273}{1283} & \frac{2858}{1283} & -\frac{2236}{1283} & \frac{1039}{1283} \\ 0 & -11 & 0 & 0 & \frac{10956}{1283} & -\frac{9735}{1283} & \frac{8041}{1283} & -\frac{3465}{1283} \\ 0 & 0 & 1 & 0 & \frac{179}{1283} & -\frac{132}{1283} & \frac{9}{1283} & \frac{40}{1283} \\ 0 & 0 & 0 & 1 & \frac{523}{1283} & -\frac{314}{1283} & \frac{313}{1283} & -\frac{177}{1283} \end{array} \right)$$

Πολλαπλασιάζουμε την δεύτερη γραμμή με $-\frac{1}{11}$ για να πάρουμε

$$\left(\begin{array}{cccc|cccc} 1 & 3 & 0 & 0 & -\frac{3273}{1283} & \frac{2858}{1283} & -\frac{2236}{1283} & \frac{1039}{1283} \\ 0 & 1 & 0 & 0 & -\frac{996}{1283} & \frac{885}{1283} & -\frac{731}{1283} & \frac{315}{1283} \\ 0 & 0 & 1 & 0 & \frac{179}{1283} & -\frac{132}{1283} & \frac{9}{1283} & \frac{40}{1283} \\ 0 & 0 & 0 & 1 & \frac{523}{1283} & -\frac{314}{1283} & \frac{313}{1283} & -\frac{177}{1283} \end{array} \right)$$

Τέλος πολλαπλασιάζουμε την δεύτερη γραμμή με -3 και την προσθέτουμε στην πρώτη

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -\frac{285}{1283} & \frac{203}{1283} & -\frac{43}{1283} & \frac{94}{1283} \\ 0 & 1 & 0 & 0 & -\frac{996}{1283} & \frac{885}{1283} & -\frac{731}{1283} & \frac{315}{1283} \\ 0 & 0 & 1 & 0 & \frac{179}{1283} & -\frac{132}{1283} & \frac{9}{1283} & \frac{40}{1283} \\ 0 & 0 & 0 & 1 & \frac{523}{1283} & -\frac{314}{1283} & \frac{313}{1283} & -\frac{177}{1283} \end{array} \right)$$

Άρα ο αντίστροφος πίνακας είναι ο

$$\left(\begin{array}{cccc} -\frac{285}{1283} & \frac{203}{1283} & -\frac{43}{1283} & \frac{94}{1283} \\ -\frac{996}{1283} & \frac{885}{1283} & -\frac{731}{1283} & \frac{315}{1283} \\ \frac{179}{1283} & -\frac{132}{1283} & \frac{9}{1283} & \frac{40}{1283} \\ \frac{523}{1283} & -\frac{314}{1283} & \frac{313}{1283} & -\frac{177}{1283} \end{array} \right)$$

Λύση της άσκησης III.5

Για να υπολογίσουμε τον αντίστροφο πίνακα θα πρέπει να δουλέψουμε με τον 2×4 πίνακα

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right)$$

Διακρίνουμε τις παρακάτω περιπτώσεις:

1. $a \neq 0$. Στην περίπτωση αυτή πολλαπλασιάζουμε την πρώτη γραμμή με $-\frac{c}{a}$ και την προσθέτουμε στην δεύτερη για να πάρουμε

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \frac{da-bc}{a} & -\frac{c}{a} & 1 \end{array} \right)$$

Αν τώρα η ποσότητα $da - bc = 0$ τότε δεν υπάρχει αντίστροφος αφού η εξίσωση που προκύπτει από την τελευταία σειρά δεν έχει λύση. Ας υποθέσουμε λοιπόν ότι $da - bc \neq 0$. Μπορούμε λοιπόν να πολλαπλασιάσουμε την τελευταία γραμμή με $\frac{a}{da-bc}$ για να πάρουμε

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array} \right)$$

Τώρα πολλαπλασιάζουμε την δεύτερη γραμμή με $-b$ και την προσθέτουμε στην πρώτη για να πάρουμε

$$\left(\begin{array}{cc|cc} a & 0 & 1 + \frac{bc}{da-bc} & -\frac{ba}{da-bc} \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array} \right)$$

Στην συνέχεια πολλαπλασιάζουμε την πρώτη γραμμή με $\frac{1}{a}$ και καταλήγουμε στον πίνακα

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{da-bc} & -\frac{b}{da-bc} \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array} \right)$$

Συνεπώς ο αντίστροφος πίνακας είναι ο

$$A^{-1} = \frac{1}{da-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

2. Σε αυτή την περίπτωση υποθέτουμε ότι $a = 0$, δηλαδή θα δουλέψουμε με τον πίνακα

$$\left(\begin{array}{cc|cc} 0 & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right)$$

Αντιμεταθέτουμε τις δύο γραμμές για να πάρουμε

$$\left(\begin{array}{cc|cc} c & d & 0 & 1 \\ 0 & b & 1 & 0 \end{array} \right)$$

Αν $b = 0$ τότε η εξίσωση που προκύπτει από την τελευταία γραμμή δεν έχει λύση. Υποθέτουμε λοιπόν ότι $b \neq 0$. Πολλαπλασιάζουμε την τελευταία γραμμή με $\frac{1}{b}$ για να πάρουμε

$$\left(\begin{array}{cc|cc} c & d & 0 & 1 \\ 0 & 1 & \frac{1}{b} & 0 \end{array} \right)$$

Πολλαπλασιάζουμε την τελευταία γραμμή με $-d$ και την προσθέτουμε στην πρώτη

$$\left(\begin{array}{cc|cc} c & 0 & -\frac{d}{b} & 1 \\ 0 & 1 & \frac{1}{b} & 0 \end{array} \right)$$

Αν $c = 0$ τότε η εξίσωση που προκύπτει από την πρώτη γραμμή δεν έχει λύση, υποθέτουμε λοιπόν ότι $c \neq 0$ και πολλαπλασιάζουμε την πρώτη γραμμή με $\frac{1}{c}$ για να πάρουμε

$$\left(\begin{array}{cc|cc} 1 & 0 & -\frac{d}{bc} & \frac{b}{bc} \\ 0 & 1 & \frac{1}{b} & 0 \end{array} \right)$$

Παρατηρούμε ότι και σε αυτή την περίπτωση η συνθήκη αντιστρεψιμότητας είναι η $0 \neq bc = ad - bc$ και ότι και πάλι ο αντίστροφος πίνακας δίνεται από τον τύπο

$$A^{-1} = \frac{1}{da - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Λύση της άσκησης III.6

Θα το αποδείξουμε με επαγωγή στο n . Αν $n = 1$ και $A = (a_{11})$ τότε

$$\det \begin{pmatrix} a_{11} & c_{11} & c_{12} & \cdots & c_{1,m} \\ 0 & b_{11} & b_{12} & \cdots & b_{1,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & b_{m,1} & b_{m,2} & \cdots & b_{m,m} \end{pmatrix} = a_{11} \det(B) = \det(A) \det(B).$$

Υποθέτουμε ότι η πρόταση ισχύει για κάθε πίνακα A ο οποίος είναι $(n-1) \times (n-1)$, και θεωρούμε τον πίνακα $A \in \mathbb{F}^{n,n}$ και τον μπλοκ πίνακα

$$\det(X) = \det \begin{pmatrix} a_{11} & \cdots & a_{1,n} & c_{11} & c_{12} & \cdots & c_{1,m} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} & c_{n,1} & c_{n,2} & \cdots & c_{n,m} \\ 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1,m} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & b_{m,1} & b_{m,2} & \cdots & b_{m,m} \end{pmatrix} = \sum_{i=1}^n (-1)^{1+i} a_{i,1} X_{i,1},$$

όπου οι ελάσσονες πίνακες είναι μπλοκ άνω τριγωνικοί με πάνω αριστερά κομμάτι τον $(n-1) \times (n-1)$ ελάσσων πίνακα A_{ij} . Από την επαγωγική υπόθεση έχουμε ότι

$$\det(X) = \sum_{i=1}^n (-1)^{1+i} a_{i,1} A_{i,1} \det(B) = \det(A) \det(B).$$

Λύση της άσκησης III.7

Έχουμε ότι $(\det A)^2 = \det(A^2) = \det(-\mathbb{I}_n) = (-1)^n$. Άρα αν ο n είναι περιττός έχουμε ότι $\det(A)^2 = -1$ το οποίο είναι αδύνατο, αφού $\det(A) \in \mathbb{R}$. Για n άρτιο δεν έχουμε τέτοια αντίφαση και παρατηρούμε ότι

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbb{I}_2.$$

Λύση της άσκησης III.8

Έστω d_n η ζητούμενη ορίζουσα. Παρατηρούμε ότι

$$d_1 = \det \begin{pmatrix} 0 & x_1 \\ x_1 & 1 \end{pmatrix} = -x_1^2.$$

Θα δείξουμε με επαγωγή ότι $d_n = -x_1^2 - x_2^2 - \dots - x_n^2$. Αναπτύσσουμε την ορίζουσα ως προς την τελευταία στήλη και έχουμε:

$$\begin{aligned} d_n &= (-1)^{n+1} x_n \det \begin{pmatrix} x_1 & 1 & 0 & \cdots & 0 \\ x_2 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ x_{n-1} & 0 & \cdots & 0 & 1 \\ x_n & 0 & \cdots & \cdots & 0 \end{pmatrix} + (-1)^{2n} d_{n-1} \\ &= (-1)^{2n+1} x_n^2 + d_{n-1}, \end{aligned}$$

οπότε το ζητούμενο προκύπτει από την επαγωγική υπόθεση.

Λύση της άσκησης III.9

Θα συμβολίζουμε τις στήλες του πίνακα A με $a^1, a^2, \dots, a^n$. Επίσης, θέτουμε $\bar{\lambda} = (\lambda, \dots, \lambda)^t$. Από την γραμμικότητα της ορίζουσας έχουμε ότι

$$\begin{aligned} D(\lambda) &= \det(\bar{\lambda} + a^1, \bar{\lambda} + a^2, \dots, \bar{\lambda} + a^n) = \det(\bar{\lambda}, a^2 + \bar{\lambda}, \dots, \bar{\lambda} + a^n) + \det(a^1, a^2 + \bar{\lambda}, \dots, \bar{\lambda} + a^n) \\ &= \det(\bar{\lambda}, a^2, \dots, \bar{\lambda} + a^n) + \det(\bar{\lambda}, \bar{\lambda}, \dots, \bar{\lambda} + a^n) + \det(a^1, a^2, \dots, \bar{\lambda} + a^n) + \det(a^1, \bar{\lambda}, \dots, \bar{\lambda} + a^n) \end{aligned}$$

Συνεχίζοντας με αυτόν τον τρόπο παίρνουμε ένα άθροισμα από ορίζουσες όπου η k -στήλη είναι ή a^k ή $\bar{\lambda}$. Όμως αν δύο στήλες είναι ίσες με $\bar{\lambda}$ τότε η ορίζουσα μηδενίζεται. Στο τελικό άθροισμα έχουμε δηλαδή

$$D(\lambda) = \sum_{v=1}^n \det(a^1, \dots, a^{v-1}, \bar{\lambda}, a^{v+1}, a^n) = A\lambda + B$$

για κάποια $A, B \in \mathbb{F}$, δηλαδή το ζητούμενο.

Για το δεύτερο ερώτημα, αναζητούμε το $D(0) = B$. Γνωρίζουμε ότι $D(\lambda_1) = A\lambda_1 + B$, $D(\lambda_2) = A\lambda_2 + B$ από όπου έχουμε ότι

$$A = \frac{D(\lambda_1) - D(\lambda_2)}{\lambda_1 - \lambda_2}$$

και συνεπώς

$$D(0) = B = D(\lambda_1) - A\lambda_1 = D(\lambda_1) - \lambda_1 \frac{D(\lambda_1) - D(\lambda_2)}{\lambda_1 - \lambda_2}.$$

Παρατηρούμε ότι $D(-a) = (z - a)^n$ και ότι $D(-b) = (z - b)^n$, συνεπώς:

$$D(0) = (z - a)^n - a \frac{(z - a)^n - (z - b)^n}{a - b} = \frac{a(z - b)^n - b(z - a)^n}{a - b}.$$

Λύση της άσκησης III.10

Θεωρούμε τους στοιχειώδεις μετασχηματισμούς γραμμών $E^1(i, j)$, $E^2(\lambda, i)$, $E^3(\lambda, i, j)$ όπως ορίστηκαν στην [III.4.1](#). Έχουμε ότι $\phi(E^2(\lambda, j)) = \lambda$. Επίσης έχουμε ότι

$$E^3(\lambda, j, i) = E^2(\lambda^{-1}, j) E^3(1, j, i) E^2(\lambda, j)$$

από όπου έχουμε ότι

$$\phi(E^3(\lambda, j, i)) = \phi(E^2(\lambda^{-1}, j)) \phi(E^3(1, j, i)) \phi(E^2(\lambda, j)) \Rightarrow \phi(E^3(\lambda, j, i)) = \phi(E^3(1, j, i)).$$

Όμως $E^3(1, j, i)^2 = E^3(2, j, i)$ συνεπώς $\phi(E^3(1, j, i))^2 = \phi(E^3(2, j, i))$, άρα $\phi(E^3(1, j, i)) = 1$. Επίσης

$$E^1(i, j) = E^2(-1, i) E^3(1, i, j) E^3(-1, j, i) E^3(1, i, j),$$

από όπου έχουμε

$$\phi(E^1(i, j)) = \phi(E^2(-1, i))\phi(E^3(1, i, j))\phi(E^3(-1, j, i))\phi(E^3(1, i, j)) = -1.$$

Δηλαδή όλοι οι στοιχειώδεις μετασχηματισμοί E έχουν $\phi(E) = \det(E)$. Συνεπώς για κάθε αντιστρέψιμο πίνακα, ο οποίος γράφεται ως γινόμενο στοιχειωδών μετασχηματισμών ισχύει ότι $\phi(E) = \det(E)$. Αν ο A δεν είναι αντιστρέψιμος τότε μετά από μια σειρά στοιχειωδών μετασχηματισμών γραμμών, δηλαδή πολλαπλασιασμών με στοιχειώδεις πίνακες μετασχηματισμών γίνεται διαγώνιος με μηδενικά στοιχεία στην διαγώνιο. Οπότε

$$\phi(A) = \det(E_1 \cdots E_s)\phi(\text{diag}(\lambda_1, \dots, \lambda_t, 0, \dots, 0)) = 0 = \det(A).$$

Λύση της άσκησης III.11

Θα αποδείξουμε μόνο την πρώτη ισότητα, η δεύτερη αποδεικνύεται παρόμοια ή ανάγεται στην πρώτη κάνοντας χρήση αναστροφής.

Θα αποδείξουμε το ζητούμενο με επαγωγή στο k . Παρατηρούμε ότι για $k = 1$ έχουμε ότι $(i_1 | j_1)$ είναι η ελάσσων ορίζουσα i_1, j_1 η οποία έχει προκύψει με διαγραφή της i_1 γραμμής και j_1 στήλης ενώ η $A \begin{pmatrix} i_1 \\ j_1 \end{pmatrix} = a_{i_1, j_1}$. Το ζητούμενο προκύπτει από το ανάπτυγμα της ορίζουσας ως προς την j_1 -στήλη.

Θεωρούμε ότι ζητούμενο είναι σωστό για $k - 1$ και θα το αποδείξουμε για k . Έχουμε

$$\begin{aligned} \det(A) &= \sum_{d_1=1}^n a_{d_1, j_1} (-1)^{d_1+j_1} \det A \begin{pmatrix} d_1 \\ j_1 \end{pmatrix} \\ &= \sum_{d_1=1}^n a_{d_1, j_1} (-1)^{d_1+j_1} \sum_{\substack{1 \leq d_2, \dots, d_k \leq n \\ d_1 \neq d_2, \dots, d_k}} (-1)^{\rho_1(d_2)+\dots+\rho_1(d_k)+j_2+\dots+j_k-(k-1)} \det A \begin{pmatrix} d_2, \dots, d_k \\ j_2, \dots, j_k \end{pmatrix} \det A \begin{pmatrix} d_1, \dots, d_k \\ j_1, \dots, j_k \end{pmatrix} \end{aligned}$$

Στον παραπάνω τύπο έχουμε εφαρμόσει την επαγωγική υπόθεση. Χρειάζεται προσοχή γιατί όταν αφαιρούμε μια γραμμή ή μια στήλη η αρίθμηση των στηλών στην νέα ορίζουσα μπορεί να διαφέρει. Με δεδομένο ότι έχουμε αφαιρέσει την πρώτη στήλη στην $A \begin{pmatrix} d_1 \\ j_1 \end{pmatrix}$ οι στήλες διαφέρουν και για αυτό έχουμε αφαιρέσει το $k - 1$ στον εκθέτη του -1 . Για τις γραμμές ορίζουμε το εξής: $\rho_1(d) = d - 1$ αν το $d_1 < d$ και $\rho_1(d) = d$ αν $d_1 > d$. Συνεχίζουμε τον υπολογισμό, όπου θέτουμε $J = j_1 + \dots + j_k$

$$\begin{aligned} &= \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \sum_{\substack{1 \leq s \leq k \\ d_1 = i_s \\ d_r = i_{r-1}, 2 \leq r \leq s \\ d_r = i_r, r > s}} (-1)^{(d_1+\rho_1(d_2)+\dots+\rho_1(d_k))+J-(k-1)} a_{d_1, j_1} \det A \begin{pmatrix} d_2, \dots, d_k \\ j_2, \dots, j_k \end{pmatrix} \det A \begin{pmatrix} d_2, \dots, d_k \\ j_2, \dots, j_k \end{pmatrix} \\ &= \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \sum_{s=1}^k (-1)^{(i_1+\dots+i_k)+J+(s-1)} a_{i_s, j_1} \det A \begin{pmatrix} i_1, \dots, i_{s-1}, i_{s+1}, \dots, i_k \\ j_2, \dots, j_k \end{pmatrix} \det A \begin{pmatrix} i_1, i_2, \dots, i_k \\ j_1, j_2, \dots, j_k \end{pmatrix} \\ &= \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \det A \begin{pmatrix} i_1, i_2, \dots, i_k \\ j_1, j_2, \dots, j_k \end{pmatrix} (-1)^{(i_1+\dots+i_k)+(j_1+\dots+j_k)} \det A \begin{pmatrix} i_1, i_2, \dots, i_k \\ j_1, j_2, \dots, j_k \end{pmatrix}. \end{aligned}$$

Λύση της άσκησης III.12

Γνωρίζουμε ότι

$$\det A = \sum_{\sigma \in S_n} \text{sgn} \sigma \prod_{1 \leq k \leq n} A_{k, \sigma(k)}.$$

Θέτουμε $[n] := \{1, \dots, n\}$ και παρατηρούμε ότι

$$\prod_{1 \leq k \leq n} (a_k + b_k) = \sum_{C \subseteq [n]} \prod_{i \in C} a_i \prod_{j \in [n]-C} b_j$$

αφού στον υπολογισμό του παραπάνω γινομένου διαλέγουμε τον όρο a_k ή τον όρο b_k και στο άθροισμα έχουμε γράψει όλους τους τρόπους με τους οποίους μπορούν να γίνουν αυτές οι επιλογές. Με αυτό τον τρόπο υπολογίζουμε:

$$\begin{aligned}\det(x\mathbb{I}_n - A) &= \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{1 \leq k \leq n} (x\delta_{k, \sigma(k)} - A_{k, \sigma(k)}) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \sum_{C \subseteq [n]} \prod_{i \in C} (-A_{i, \sigma(i)}) \prod_{j \in [n] - C} t\delta_{j, \sigma(j)} \\ &= \sum_{C \subseteq [n]} (-1)^{|C|} \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{i \in C} A_{i, \sigma(i)} \prod_{j \in [n] - C} t\delta_{j, \sigma(j)}.\end{aligned}$$

Για ένα σταθερό $C \subseteq [n]$ και $\sigma \in S_n$ το τελευταίο γινόμενο $\prod_{j \in [n] - C} x\delta_{j, \sigma(j)}$ μηδενίζεται εκτός αν τα στοιχεία του $[n] - C$ κρατούνται σταθερά από το σ , στην περίπτωση αυτή το γινόμενο είναι απλά το $x^{n-|C|}$. Δηλαδή χρειάζεται να θεωρήσουμε μόνο τις συνεισφορές των μεταθέσεων των C στο άθροισμά μας. Θεωρούμε ένα στοιχείο $\sigma \in S_n$ που σταθεροποιεί το $[n] - C$ ως ένα στοιχείο του S_C . Το πρόσημο μιας τέτοιας μετάθεσης του S_C παραμένει το ίδιο, όπως βλέπουμε αναλύοντας την μετάθεση σ ως γινόμενο αντιμεταθέσεων. Καταλήγουμε στο

$$\begin{aligned}\sum_{C \subseteq [n]} (-1)^{|C|} \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{i \in C} A_{i, \sigma(i)} \prod_{j \in [n] - C} t\delta_{j, \sigma(j)} &= \sum_{C \subseteq [n]} (-1)^{|C|} \sum_{\sigma \in S_C} \operatorname{sgn} \sigma \prod_{i \in C} A_{i, \sigma(i)} t^{n-|C|} \\ &= \sum_{C \subseteq [n]} (-1)^{|C|} t^{n-|C|} \sum_{\sigma \in S_C} \operatorname{sgn} \sigma \prod_{i \in C} A_{i, \sigma(i)}.\end{aligned}$$

Ο όρος

$$\sum_{\sigma \in S_C} \operatorname{sgn} \sigma \prod_{i \in C} A_{i, \sigma(i)}$$

είναι η ορίζουσα του πίνακα $A[C]$ οπότε έχουμε

$$\begin{aligned}\sum_{C \subseteq [n]} (-1)^{|C|} x^{n-|C|} \sum_{\sigma \in S_C} \operatorname{sgn} \sigma \prod_{i \in C} A_{i, \sigma(i)} &= \sum_{C \subseteq [n]} (-1)^{|C|} x^{n-|C|} \det(A[C]) \\ &= \sum_{0 \leq k \leq n} \sum_{\substack{C \subseteq [n] \\ |C|=k}} (-1)^k t^{n-k} \det(A[C]) \\ &= \sum_{0 \leq k \leq n} x^{n-k} \left((-1)^k \sum_{\substack{C \subseteq [n] \\ |C|=k}} \det(A[C]) \right) \\ &= \sum_{0 \leq k \leq n} x^k \left((-1)^{n-k} \sum_{\substack{C \subseteq [n] \\ |C|=n-k}} \det(A[C]) \right).\end{aligned}$$

Λύση της άσκησης III.13

Θεωρούμε τους μπλοκ πίνακες

$$\mathbb{F}^{n+m, n+m} \ni C = \begin{pmatrix} x\mathbb{I}_m & A \\ B & \mathbb{I}_n \end{pmatrix} \quad D = \begin{pmatrix} \mathbb{I}_m & \mathbf{0}_{m,n} \\ -B & x\mathbb{I}_n \end{pmatrix} \in \mathbb{F}^{n+m, n+m}.$$

και υπολογίζουμε ότι

$$CD = \begin{pmatrix} x\mathbb{I}_m - AB & xA \\ \mathbf{0}_{n,m} & x\mathbb{I}_n \end{pmatrix} \quad DC = \begin{pmatrix} x\mathbb{I}_m & A \\ \mathbf{0}_{n,m} & -BA + x\mathbb{I}_n \end{pmatrix}$$

Κάνοντας χρήση του $\det(CD) = \det(DC)$ και της άσκησης III.6 καταλήγουμε στο ζητούμενο.

Λύση της άσκησης III.14

Συγκρίνουμε τους συντελεστές της ισότητας της άσκησης III.13 και για τον συντελεστή του x^{n-m} το αριστερό μέρος το υπολογίζουμε με την βοήθεια της άσκησης III.12 ενώ το αριστερό είναι η ορίζουσα $\det(AB)$, δηλαδή

$$\det(AB) = \sum_{S \in \binom{[n]}{n}} \det(BA[S]) = \sum_{S \in \binom{[n]}{n}} \det(B_{S,m}) \det(A_{m,S}),$$

δηλαδή το ζητούμενο.

Λύση της άσκησης IV.1

Θεωρούμε δύο συγκλίνουσες ακολουθίες $(a_n), (b_n)$. Είναι γνωστό ότι το άθροισμα τους θα είναι συγκλίνουσα ακολουθία και επίσης η (ℓa_n) είναι συγκλίνουσα για κάθε $\ell \in \mathbb{R}$. Το ζητούμενο προκύπτει από την πρόταση IV.1.4

Λύση της άσκησης IV.2

Υπενθυμίζουμε ότι μία ακολουθία είναι τελικά σταθερή αν και μόνο αν υπάρχει $n_0 \in \mathbb{N}$ ώστε $n \geq n_0$ να ισχύει $a_n = a_{n_0}$. Θεωρούμε δύο τελικά σταθερές ακολουθίες $(a_n), (b_n)$ δηλαδή δύο ακολουθίες $(a_n), (b_n)$, ώστε για κατάλληλους δείκτες n_0, m_0 να έχουμε

$$a_n = a_{n_0} \text{ για } n \geq n_0 \text{ και } b_n = b_{m_0} \text{ για } n \geq m_0.$$

Η ακολουθία $(a_n + b_n)$ είναι τελικά σταθερή μετά τον δείκτη $\max(n_0, m_0)$, και η ακολουθία (ℓa_n) είναι προφανώς τελικά σταθερή.

Λύση της άσκησης IV.3

Παρατηρήστε ότι το άθροισμα δύο ακολουθιών που συγκλίνουν στο 0 συγκλίνει στο 0, και αν $a_n \rightarrow 0$ τότε και $\ell a_n \rightarrow 0$ για κάθε $\ell \in \mathbb{R}$.

Λύση της άσκησης IV.4

Το W είναι μη κενό, οπότε αν $x \in W$, τότε και $-x \in W$. Άρα $0 = x + (-x) \in W$.

Λύση της άσκησης IV.5

Γνωρίζουμε ότι αφού το $\{v_1, \dots, v_n, v\}$ είναι γραμμικά εξαρτημένο ένα από τα διανύσματα $v_1, \dots, v_n, v$ γράφεται ως γραμμικός συνδυασμός των υπολοίπων. Θα δείξουμε ότι το v είναι ένα τέτοιο. Πράγματι έστω μια μη-τετριμμένη σχέση εξάρτησης

$$\lambda_1 v_1 + \dots + \lambda_n v_n + \lambda v = 0$$

όπου δεν είναι όλα τα $\lambda_1, \dots, \lambda_n, \lambda$ ίσα με 0. Αν $\lambda = 0$ τότε η παραπάνω σχέση γίνεται μια μη τετριμμένη σχέση εξάρτησης ανάμεσα στα γραμμικά ανεξάρτητα $v_1, \dots, v_n$, άτοπο.

Άρα $\lambda \neq 0$ και

$$v = - \sum_{\mu=1}^n \frac{\lambda_\mu}{\lambda} v_\mu.$$

Λύση της άσκησης IV.6

Θεωρούμε την εξίσωση $x_1 e_1 + x_2 e_2 + x_3 e_3 = 0$. Αρκεί να αποδείξουμε ότι δέχεται σαν μοναδική λύση το $x_1 = x_2 = x_3 = 0$. Πράγματι,

$$x_1 e_1 + x_2 e_2 + x_3 e_3 = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1) = (x_1, x_2, x_3),$$

άρα $x_1 e_1 + x_2 e_2 + x_3 e_3 = (0, 0, 0) \Leftrightarrow x_1 = x_2 = x_3 = 0$.

Λύση της άσκησης IV.7

Θεωρούμε την εξίσωση $x_1v_1 + x_2v_2 + x_3v_3 = 0$. Αρκεί να αποδείξουμε ότι δέχεται σαν μοναδική λύση το $x_1 = x_2 = x_3 = 0$. Πράγματι,

$$x_1v_1 + x_2v_2 + x_3v_3 = (x_1 + x_3, x_1 + 3x_2, 2x_1 + x_2).$$

Το τελευταίο διάνυσμα είναι το μηδενικό αν και μόνο αν

$$\begin{aligned} x_1 + x_3 &= 0 & x_1 &= -x_3 \\ x_1 + 3x_2 &= 0 & \Leftrightarrow x_1 &= -3x_2 \\ 2x_1 + x_2 &= 0 & x_1 &= -\frac{1}{2}x_2 \end{aligned}$$

Εύκολα παρατηρεί κανείς στο παραπάνω σύστημα ότι μοναδική λύση είναι η $x_1 = x_2 = x_3 = 0$.

Λύση της άσκησης IV.8

Υπολογίζουμε ότι

$$x_1w_1 + x_2w_2 + x_3w_3 = (x_1 + x_2 + 2x_3, x_1 + 3x_2 + 4x_3, 2x_1 + 4x_2 + 6x_3, 2x_1 + 5x_2 + 7x_3).$$

Το παραπάνω διάνυσμα είναι μηδενικό αν και μόνο αν

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 0 \\ x_1 + 3x_2 + 4x_3 &= 0 \\ 2x_1 + 4x_2 + 6x_3 &= 0 \\ 2x_1 + 5x_2 + 7x_3 &= 0 \end{aligned}$$

Όμως το παραπάνω σύστημα εκτός της μηδενικής, έχει σαν λύση και την $x_1 = 1, x_2 = 1, x_3 = 2$, άρα τα διανύσματα είναι γραμμικά εξαρτημένα.

Λύση της άσκησης IV.9

Αφού τα διανύσματα $v_1, v_2, \dots, v_n$ παράγουν τον χώρο V , το διάνυσμα $v \in V$, μπορεί να γραφεί ως γραμμικός συνδιασμός των $v_1, \dots, v_n$, άρα τα $v_1, \dots, v_n, v$ είναι γραμμικά εξαρτημένα.

Λύση της άσκησης IV.10

Έστω $v \in V$, και έστω $v = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i$, δύο διαφορετικές εκφράσεις του v ως γραμμικός συνδιασμός των v_i . Αφαιρώντας κατά μέλη έχουμε ότι

$$\sum_{i=1}^n (\lambda_i - \mu_i) v_i = 0,$$

οπότε η γραμμική ανεξαρτησία των v_i επιβάλλει $\lambda_i = \mu_i$, δηλαδή η γραφή είναι μοναδική.

Λύση της άσκησης IV.12

Λάθος, αφού για παράδειγμα στον γραμμικό χώρο $\mathbb{R}^n$, τα διανύσματα

$$(1, 0, \dots, 0), (2, 0, \dots, 0), \dots, (n, 0, \dots, 0)$$

Λύση της άσκησης IV.13

Λάθος, αφού από τον ορισμό της βάσης, αυτή αποτελείται από n το πλήθος γραμμικά ανεξάρτητα διανύσματα.

Λύση της άσκησης IV.14

Σωστό. Αν σε ένα χώρο V έχουμε m γραμμικά ανεξάρτητα διανύσματα με $m > n$ τότε αυτά τα διανύσματα θα παράγουν ένα υπόχωρο W , με $W \subseteq V$, και συνεπώς $\dim W \leq \dim V = n$. Όμως η διάσταση του W είναι m , αφού παράγεται από m το πλήθος γραμμικά ανεξάρτητα διανύσματα, άτοπο.

Λύση της άσκησης IV.15

Ορίζουμε πρόσθεση και πολλαπλασιασμό κατά συντεταγμένες, δηλαδή

$$+ : \begin{cases} (V \times W) \times (V \times W) & \rightarrow & V \times W \\ ((v_1, w_1), (v_2, w_2)) & \mapsto & (v_1 + v_2, w_1 + w_2) \end{cases}$$

και

$$\cdot : \begin{cases} \mathbb{R} \times (V \times W) & \rightarrow & V \times W \\ (\lambda, (v, w)) & \mapsto & (\lambda v, \lambda w). \end{cases}$$

Οι ιδιότητες των πράξεων που εξασφαλίζουν ότι ο $V \times W$, προκύπτουν από τις ιδιότητες των V, W . Για παράδειγμα η προσεταιριστική ιδιότητα ανάγεται στην προσεταιριστική ιδιότητα των χώρων V, W :

$$\begin{aligned} ((v_1, w_1) + (v_2, w_3)) + (v_3, w_3) &= (v_1 + v_2, w_1 + w_2) + (v_3, w_3) = \\ ((v_1 + v_2) + v_3, (w_1 + w_2) + w_3) &= (v_1 + (v_2 + v_3), w_1 + (w_2 + w_3)) = \\ = (v_1, w_1) + (v_2 + v_3, w_2 + w_3) &= (v_1, w_1) + ((v_2, w_2) + (v_3, w_3)). \end{aligned}$$

Επιλέγουμε μία βάση $v_1, \dots, v_n$ του V και μια βάση $w_1, \dots, w_m$ του W . Τα στοιχεία (v_i, w_j) $i = 1, \dots, n, j = 1, \dots, m$ αποτελούν βάση του $V \times W$. Πράγματι, κάθε στοιχείο (v, w) γράφεται ως

$$\left(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^m \mu_j w_j \right) = \sum_{i=1}^n \sum_{j=1}^m \lambda_i \mu_j (v_i, w_j).$$

Επίσης τα παραπάνω στοιχεία είναι γραμμικά ανεξάρτητα, αφού αν

$$\sum_{i=1}^n \sum_{j=1}^m \lambda_i \mu_j (v_i, w_j) = (0, 0),$$

τότε

$$\sum_{i=1}^n \lambda_i v_i = \sum_{j=1}^m \mu_j w_j = 0,$$

και $\lambda_i = \mu_j = 0$ για όλα τα i, j , λόγω της γραμμικής ανεξαρτησίας των v_i, w_j .

Λύση της άσκησης IV.16

Εξ ορισμού ο χώρος που παράγουν οι συναρτήσεις x^i είναι διανυσματικός υπόχωρος του χώρου όλων των συναρτήσεων. Για να αποδείξουμε ότι οι συναρτήσεις $\{1, x, x^2, \dots, x^n\}$ είναι γραμμικά ανεξάρτητες θεωρούμε τον γραμμικό συνδιασμό

$$\sum_{i=0}^n \lambda_i x^i = 0.$$

Παρατηρούμε ότι το αριστερό μέρος της παραπάνω εξίσωσης αν κάποιο $\lambda_i \neq 0$ δεν μπορεί να είναι ταυτοτικά ίσο με μηδέν; αποτελεί μία πολυωνυμική συνάρτηση που έχει πεπερασμένες το πλήθος ρίζες.

Λύση της άσκησης IV.17

Αρκεί να αποδείξουμε ότι το άθροισμα δύο συνεχών συναρτήσεων είναι συνεχής συνάρτηση και το γινόμενο συνεχούς συνάρτησης επί σταθεράς είναι συνεχής συνάρτηση. Θα κάνουμε το πρώτο που είναι και δυσκολότερο. Έστω f, g συνεχείς συναρτήσεις, δηλαδή για κάθε συγκλίνουσα ακολουθία $a_n \rightarrow \ell \in \mathbb{R}$, έχουμε

$$\lim f(a_n) = f(\lim a_n) \text{ και } \lim g(a_n) = g(\lim a_n).$$

Παρατηρούμε ότι

$$\begin{aligned} \lim(f+g)(a_n) &= \lim(f(a_n) + g(a_n)) = \lim f(a_n) + \lim g(a_n) = \\ &= f(\lim a_n) + g(\lim a_n) = (f+g)(\lim a_n), \end{aligned}$$

από όπου προκύπτει το ζητούμενο.

Λύση της άσκησης V.1

Αν η συνάρτηση είναι γραμμική τότε

$$(\lambda v + \mu w) = T(\lambda v) + T(\mu w) = \lambda T(v) + \mu T(w).$$

Αντιστρόφως αν ισχύει η (V.1) τότε για $\lambda = \mu = 1$, έχουμε

$$T(v+w) = T(v) + T(w),$$

ενώ για $\mu = 0$ έχουμε

$$(\lambda v) = \lambda T(v),$$

δηλαδή η T είναι γραμμική.

Λύση της άσκησης V.2

Θεωρούμε μία γραμμική απεικόνιση $T: \mathbb{F} \rightarrow \mathbb{F}$. Η τιμή της στο $x \in \mathbb{F}$, δίνεται από

$$T(x) = T(x \cdot 1) = xT(1),$$

όπου $T(1) \in \mathbb{F}$. Άρα οι γραμμικές συναρτήσεις είναι οι συναρτήσεις τις μορφής

$$x \mapsto ax,$$

όπου το a διατρέχει το $\mathbb{F}$.

Λύση της άσκησης V.3

Έστω $x \in V$, έχουμε $T(0) = T(x-x) = T(x) - T(x) = 0$.

Λύση της άσκησης V.4

Θεωρούμε μια βάση $v_1, \dots, v_n$ του V . Αφού η f είναι μονομορφισμός τα $f(v_1), \dots, f(v_n)$ είναι γραμμικά ανεξάρτητα, σύμφωνα με την παρατήρηση V.1.18. Οπότε, σύμφωνα με την πρόταση V.1.25 υπάρχει $g: W \rightarrow V$ με $g \circ f(v_i) = v_i$. Άρα στα στοιχεία της βάσης έχουμε $g \circ f(v_i) = \text{Id}_V(v_i)$ $g \circ f = \text{Id}_V$.

Αντιστρόφως αν υπάρχει $g: W \rightarrow V$ συνάρτηση με $g \circ f = \text{Id}_W$ και $v_1, v_2 \in V$ με $f(v_1) = f(v_2)$. Τότε $v_1 = g(f(v_1)) = g(f(v_2)) = v_2$ συνεπώς η f , είναι 1-1.

Λύση της άσκησης V.5

Έστω $w_1, \dots, w_n$ μια βάση στο W . Αφού η f είναι επί υπάρχουν στοιχεία $v_1, \dots, v_n \in V$ με $f(v_1) = w_1, \dots, f(v_n) = w_n$. Σύμφωνα με την πρόταση **V.1.25** υπάρχει γραμμική συνάρτηση $g : W \rightarrow V$ με $g(w_i) = v_i$, αφού $w_1, \dots, w_n$ είναι γραμμικά ανεξάρτητα. Παρατηρούμε ότι $f \circ g(w_i) = w_i = \text{Id}_W(w_i)$, άρα $f \circ g = \text{Id}_W$.

Αντιστρόφως αν υπάρχει συνάρτηση $g : W \rightarrow V$ με $f \circ g = \text{Id}_W$, τότε για κάθε $w \in W$ έχουμε ότι $w = \text{Id}_W(w) = f(g(w))$, με $g(w) \in V$, άρα η f είναι επί.

Λύση της άσκησης V.6

Για $n = 1$, αυτό που ζητούμε είναι προφανές. Ας υποθέσουμε ότι ισχύει για $n = k$ δηλαδή ότι ισχύει $A^k = Q^{-1}B^kQ$, και ας υπολογίσουμε

$$A^{k+1} = A^k A = (Q^{-1}B^kQ)Q^{-1}BQ = Q^{-1}B^k(QQ^{-1})BQ = Q^{-1}B^{k+1}Q,$$

δηλαδή το ζητούμενο.

Λύση της άσκησης V.7

Για την ανακλαστική σχέση, παρατηρούμε ότι κάθε πίνακας είναι συζυγής με τον ευατό του. Πράγματι, αρκεί να πάρουμε σαν Q τον $\mathbb{I}_n$. Για να αποδείξουμε την ανακλαστική σχέση θα πρέπει να δείξουμε ότι αν ο A είναι συζυγής με τον B τότε και ο B είναι συζυγής με τον A . Πράγματι, έστω ότι υπάρχει αντιστρέψιμος πίνακας Q , ώστε $A = Q^{-1}BQ$. Πολλαπλασιάζοντας την προηγούμενη σχέση από αριστερά με Q και από δεξιά με Q^{-1} έχουμε

$$QAQ^{-1} = Q(Q^{-1}BQ)Q^{-1} = QQ^{-1}BQQ^{-1} = B,$$

δηλαδή το ζητούμενο. Τέλος για την μεταβατική ιδιότητα, θα πρέπει να αποδείξουμε ότι αν ο πίνακας A είναι συζυγής με τον B και αν ο B είναι συζυγής με τον C , τότε και ο A είναι συζυγής με τον C . Υποθέτουμε ότι υπάρχουν Q, R , ώστε $A = Q^{-1}BQ$ και $B = R^{-1}CR$. Αντικαθιστώντας από την μία σχέση στην άλλη έχουμε ότι

$$A = Q^{-1}R^{-1}CRQ = (RQ)^{-1}CRQ$$

(δες και άσκηση ??) άρα ο A είναι συζυγής με τον C .

Λύση της άσκησης V.8

Έστω A ο ζητούμενος πίνακας με στοιχεία a_{ij} . Θεωρούμε τους πίνακες $B(i_0 j_0)$ με στοιχεία

$$b_{ij} = \begin{cases} 0 & \text{αν } (i, j) \neq (i_0, j_0) \\ 1 & \text{αν } (i, j) = (i_0, j_0) \end{cases}$$

Θα αποδείξουμε πρώτα ότι $a_{ij} = 0$ αν $i \neq j$. Θεωρούμε $i_0 \neq j_0$ και υπολογίζουμε τους πίνακες

$$\begin{pmatrix} 0 & \cdots & 0 & a_{1i_0} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_{2i_0} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{ni_0} & 0 & \cdots & 0 \end{pmatrix} = (i_0 j_0) =$$

$$(i_0 j_0)A = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ a_{1j_0} & \cdots & a_{nj_0} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

Από όπου προκύπτει ότι $a_{ij} \neq 0$ αν και μόνο αν $i = j$. Θεωρούμε τώρα τον ταυτοτικό πίνακα και αντιμετωπίζουμε την i -γραμμή με την j -γραμμή, ονομάζουμε τον πίνακα που προκύπτει $S(ij)$. Αν ο Ab διαγώνιος πίνακας με $\lambda_1, \dots, \lambda_n$ στην διαγώνιο, δηλαδή, τότε ο πίνακας SAS^{-1} είναι ένας πίνακας που το λ_i στοιχείο έχει αλλάξει θέση με το λ_j (γιατί?) και επιπλέον είναι ίσος με τον A . Άρα όλα τα στοιχεία στην διαγώνιο είναι ίσα, δηλαδή το ζητούμενο.

Λύση της άσκησης V.9

Παρατηρούμε ότι

$$A^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}.$$

Θα αποδείξουμε με επαγωγή ότι

$$A^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}.$$

Για $n = 1$ ισχύει, υποθέτουμε ότι ισχύει για $n = k$ και βάσει της υπόθεσης αυτής θα αποδείξουμε τον τύπο για $n = k + 1$.

$$A^{k+1} = A^k A = A^n = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix} = \begin{pmatrix} \lambda^{k+1} & (k+1)\lambda^k \\ 0 & \lambda^{k+1} \end{pmatrix},$$

δηλαδή το ζητούμενο.

Λύση της άσκησης V.10

Έστω A ο $n \times n$ πίνακας που αντιστοιχεί στην γραμμική απεικόνιση. Παρατηρούμε ότι για να μην είναι η γραμμική συνάρτηση ένα προς ένα πρέπει και αρκεί το σύστημα $Ax = 0$ να επιδέχεται εκτός της τριμμένης ($x = 0$) και άλλες λύσεις. Εκτελώντας την μέθοδο της απαλειφής του Γκαουβ παρατηρούμε ότι ο πίνακας A μπορεί να μετασχηματιστεί με τη χρήση στοιχειωδών μετασχηματισμών γραμμών σε ένα πίνακα A' της μορφής

$$\begin{pmatrix} d_1 & * & * & \cdots & * \\ 0 & d_2 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & * & * \\ 0 & & 0 & d_{n-1} & * \\ 0 & 0 & \cdots & 0 & d_n \end{pmatrix}$$

Το σύστημα έχει μοναδική λύση την $x = 0$ αν και μόνο αν $d_i \neq 0$ για κάθε $i = 1, \dots, n$. Σε αυτή την περίπτωση όποιο και να είναι το διάνυσμα στήλη b που θα τοποθετήσουμε στον επαυξημένο πίνακα προκειμένου να λύσουμε το μη ομογενές σύστημα $Ax = b$ μπορούμε να δούμε ότι θα έχουμε λύση, δηλαδή η γραμμική συνάρτηση είναι επί.

Αντιστρόφως αν η συνάρτηση είναι επί τότε δεν μπορεί το d_n να είναι μηδέν γιατί τότε το σύστημα

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

δεν θα είχε λύση για $b_n \neq 0$. Παρατηρούμε τώρα ότι ούτε το d_{n-1} θα μπορούσε να είναι 0 αφού τότε η εξίσωση

$$d_{n-1}x_{n-1} = b_{n-1} - x_{n-1}a'_{n-1,n}$$

δεν θα είχε λύση αν διαλέγαμε το b_{n-1} ώστε $b_{n-1} - x_{n-1}a'_{n-1,n}$. Συνεχίζοντας αναδρομικά βλέπουμε ότι όλα τα d_i είναι διαφορετικά του μηδενός και άρα το ομογενές σύστημα $Ax = 0$ έχει μοναδική λύση την τριμμένη, δηλαδή η γραμμική συνάρτηση είναι ένα προς ένα.

Λύση της άσκησης V.11

Υποθέτουμε ότι ο $\det(A) = 0$. Τότε υπάρχει ένα μη-μηδενικό $x = (x_1, \dots, x_n)^t \in \mathbb{C}^n$ ώστε $Ax = 0$. Έστω i_0 ο δείκτης ώστε

$$|x_{i_0}| = \max\{|x_1|, \dots, |x_n|\}.$$

Συνεπώς έχουμε

$$a_{i_0,1}x_1 + a_{i_0,2}x_2 + \dots + a_{i_0,i_0}x_{i_0} + a_{i_0,n}x_n = 0$$

το οποίο δίνει

$$|a_{i_0,i_0}| |x_{i_0}| = \left| \sum_{\substack{i=1 \\ i \neq i_0}}^n a_{i_0,i} x_i \right| \leq \sum_{\substack{i=1 \\ i \neq i_0}}^n |a_{i_0,i}| \cdot |x_i| \leq |x_{i_0}| \sum_{\substack{i=1 \\ i \neq i_0}}^n |a_{i_0,i}|$$

δηλαδή

$$|a_{i_0,i_0}| \leq \sum_{\substack{i=1 \\ i \neq i_0}}^n |a_{i_0,i}|$$

το οποίο είναι άτοπο.

Λύση της άσκησης VI.1

Υπολογίζουμε

$$\text{Ch}_A(x) = \det(A - x\mathbb{I}_n) = \begin{vmatrix} a_{11} - x & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22} - x & \dots & \dots & a_{2n} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & a_{nn} - x \end{vmatrix} = \prod_{i=1}^n (a_{ii} - x).$$

Λύση της άσκησης VI.2

Υπολογίζουμε ότι

$$\text{Ch}_A(x) = \det(A - x\mathbb{I}_n) = \det(A - x\mathbb{I}_n)^t = \det(A^t - x\mathbb{I}_n) = \text{Ch}_{A^t}(x).$$

Λύση της άσκησης VI.3

Ξεκινάμε υπολογίζοντας το χαρακτηριστικό πολυώνυμο του A ,

$$\begin{aligned} \text{Ch}_A(x) &= \det(A - x\mathbb{I}_n) = \det \begin{pmatrix} 1-x & 1 & \dots & 1 \\ 1 & 1-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 1-x \end{pmatrix} = (n-x) \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 1-x \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & -x & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -x \end{pmatrix} = (-1)^n (x-n)x^{n-1}. \end{aligned}$$

Στον παραπάνω υπολογισμό αρχικά προσθέσαμε όλες τις στήλες στην πρώτη, ενώ στην στην συνέχεια αφαιρέσαμε την πρώτη γραμμή από τις υπόλοιπες. Παρατηρούμε ότι έχουμε το n ως ιδιοτιμή με αλγεβρική πολλαπλότητα $\mu(n) = 1$ και συνεπώς με γεωμετρική πολλαπλότητα $\rho(n) = 1$, αφού $1 \leq \rho(n) \leq \mu(n) = 1$. Από την άλλη έχουμε την ιδιοτιμή 0 με αλγεβρική πολλαπλότητα $\mu(0) = n-1$. Η γεωμετρική πολλαπλότητα $\rho(0) = \dim \text{Ker} A = n-1$ αφού η τάξη του πίνακα A

είναι $r(A) = 1$. Άρα ο πίνακας είναι διαγωνοποιήσιμος αφού η αλγεβρική πολλαπλότητα κάθε ιδιοτιμής ταυτίζεται με την γεωμετρική της πολλαπλότητα.

Υπολογίζουμε τώρα τον ιδιόχωρο της ιδιοτιμής n . Θα πρέπει να λύσουμε το ομογενές σύστημα

$$\begin{pmatrix} 1-n & 1 & \cdots & 1 \\ 1 & 1-n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1-n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

το οποίο έχει μονοδιάστατο χώρο λύσεων που παράγεται από το $(1, \dots, 1)^t$.

Ο ιδιόχωρος της ιδιοτιμής 0 είναι ίσος με τον πυρήνα της A ο οποίος παράγεται από τα διανύσματα

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix}, \dots, v_{n-1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

Υπολογίζουμε τον πίνακα

$$Q = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

και έχουμε

$$Q^{-1}AQ = \text{diag}(n, 0, \dots, 0).$$

Λύση της άσκησης VI.4

Ξεκινάμε υπολογίζοντας το χαρακτηριστικό πολυώνυμο του πίνακα A

$$\begin{aligned} \text{Ch}_A(x) &= \det \begin{pmatrix} a-x & b & \cdots & b \\ b & a-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a-x \end{pmatrix} = \det \begin{pmatrix} a-x+(n-1)b & b & \cdots & b \\ a-x+(n-1)b & a-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ a-x+(n-1)b & \cdots & b & a-x \end{pmatrix} \\ &= (a+(n-1)b-x) \det \begin{pmatrix} 1 & b & \cdots & b \\ 1 & a-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ 1 & \cdots & b & a-x \end{pmatrix} \\ &= (a+(n-1)b-x) \det \begin{pmatrix} 1 & b & \cdots & \cdots & b \\ 0 & a-x-b & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & a-x-b \end{pmatrix} \\ &= (a+(n-1)b-x)(a-b-x)^{n-1}. \end{aligned}$$

Η ιδιοτιμή $(n-1)b+a$ έχει αλγεβρική πολλαπλότητα και γεωμετρική πολλαπλότητα ίση με 1.

Υπολογίζουμε τον ιδιόχωρο $\text{Ker}(A - (a + (n-1)b)\mathbb{I}_n)$ ως λύση του ομογενούς συστήματος,

$$\begin{pmatrix} -(n-1)b & b & \cdots & b \\ b & -(n-1)b & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & -(n-1)b \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ο οποίος έχει μονοδιάστατη λύση $(1, \dots, 1)^t$.

Στην συνέχεια υπολογίζουμε τον ιδιόχωρο $\text{Ker}(A - (b-a)\mathbb{I}_n)$ της ιδιοτιμής $a-b$ ως λύση του ομογενούς συστήματος

$$\begin{pmatrix} b & b & \cdots & b \\ b & b & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & b \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

το οποίο έχει λύσεις:

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, v_{n-1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

Υπολογίζουμε τον πίνακα

$$Q = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

και έχουμε

$$Q^{-1}AQ = \text{diag}(n, 0, \dots, 0).$$

Λύση της άσκησης VI.5

Αν $a = b$ το χαρακτηριστικό πολυώνυμο έχει υπολογιστεί στην άσκηση VI.4 και είναι ίσο με $((n-1)b-x)(-b-x)^{n-1}$. Υποθέτουμε ότι $a \neq b$ και υπολογίζουμε

$$\text{Ch}_A(x) = \det \begin{pmatrix} -x & a & \cdots & a \\ b & -x & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ b & \cdots & b & -x \end{pmatrix} = \frac{a(-x-b)^n - b(-x-a)^n}{a-b}$$

σύμφωνα με την άσκηση III.9.

Λύση της άσκησης VI.6

Το ζητούμενο προκύπτει από την άσκηση III.13.

Λύση της άσκησης VI.7

Το ζητούμενο προκύπτει άμεσα από την άσκηση VI.6.

Εναλλακτικά, θα μπορούσαμε να επιχειρηματολογήσουμε ως εξής: Αν A είναι αντιστρέψιμος τότε

$$A^{-1}(AB)A = BA,$$

δηλαδή οι πίνακες AB και BA είναι όμοιοι και έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Αν ο A δεν είναι αντιστρέψιμος τότε υπάρχει μια ακολουθία αντιστρέψιμων πινάκων A_n ώστε $A_n \rightarrow A$, δείτε το θεώρημα III.8.3. Ισχύει ότι

$$\begin{aligned} \text{Ch}_{AB}(x) &= \det\left(\lim_{n \rightarrow \infty} A_n B - xI_n\right) = \lim_{n \rightarrow \infty} \det(A_n B - xI_n) = \lim_{n \rightarrow \infty} \det(BA_n - xI_n) \\ &= \det\left(B \lim_{n \rightarrow \infty} A_n - xI_n\right) = \det(BA - xI_n) = \text{Ch}_{BA}(x). \end{aligned}$$

Λύση της άσκησης VI.8

Έστω λ μια ιδιοτιμή του T , οπότε θα πρέπει να υπάρχει μη-μηδενική συνάρτηση f με

$$\int_0^x f(t)dt = \lambda f(x). \quad (\text{VI.6})$$

Η παραπάνω ισότητα εξασφαλίζει ότι το ιδιοδιάνυσμα f είναι παραγωγίσιμο οπότε έχουμε

$$f(x) = \lambda f'(x) \rightarrow f(x) = ce^{\lambda x}.$$

Όμως από την (??) έχουμε ότι $c = 0$.

Λύση της άσκησης VI.9

Ψάχνουμε για συναρτήσεις τέτοιες ώστε $f'(x) = \lambda f(x)$. Αυτές είναι οι συναρτήσεις της μορφής $ce^{\lambda x}$. Δηλαδή κάθε πραγματικός αριθμός λ είναι ιδιοτιμή της D με μονοδιάστατο ιδιόχωρο $\langle e^{\lambda x} \rangle$.

Λύση της άσκησης VI.10

Παρατηρούμε ότι ο πίνακας

$$A = \begin{pmatrix} 5 & 2 \\ -3 & -1 \end{pmatrix}$$

έχει ιδιοτιμές $2 \pm \sqrt{3}$ και συνεπώς είναι διαγωνοποιήσιμος. Η ζητούμενη έκφραση είναι το $\text{tr}(A^n)$ υπολογισμένη είτε ως δύναμη του πίνακα A συνεπώς είναι μια ακέραια τιμή είτε ως δύναμη του διαγώνιου πίνακα.

Λύση της άσκησης VI.11

Αφού ο A είναι αντιστρέψιμος καμία ιδιοτιμή του δεν είναι μηδενική. Πράγματι αν υπήρχε ιδιοτιμή $\lambda = 0$ τότε ο ιδιόχωρος της E_0 θα ήταν μη-τετριμμένος. Όμως ο ιδιόχωρος της μηδενικής ιδιοτιμής ταυτίζεται με τον πυρήνα του A , άτοπο. Έστω λ μια ιδιοτιμή του πίνακα A . Έχουμε ότι υπάρχει ιδιοδιάνυσμα $v \neq 0$, ώστε

$$Av = \lambda v \text{ συνεπώς } v = \lambda A^{-1}v \text{ δηλαδή } A^{-1}v = \lambda^{-1}v.$$

Λύση της άσκησης VI.12

Αφού οι πίνακες A, A^{-1} είναι όμοιοι έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Συνεπώς αν $\lambda_1, \dots, \lambda_n$ είναι οι ιδιοτιμές του A τότε κάθε ιδιοτιμή λ_i είναι ίση με μια ιδιοτιμή λ_j^{-1} . Δεν είναι δυνατόν ένα πεπερασμένο σύνολο με περιττό αριθμό στοιχείων να το χωρίσουμε σε ζευγάρια οπότε για τουλάχιστον μία ιδιοτιμή ισχύει $\lambda = \lambda^{-1}$, συνεπώς $\lambda = \pm 1$.

Λύση της άσκησης VI.13

Ένας πίνακας τάξης 1 έχει πυρήνα τάξης $n-1$ και αυτός είναι ο ιδιόχωρος της ιδιοτιμής 0. Επίσης κάθε πίνακας είναι όμοιος με ένα άνω τριγωνικό πίνακα ο οποίος έχει στην διαγώνιο $n-1$ μηδενικά στοιχεία και ένα στοιχείο λ . Αφού όμοιοι πίνακες έχουν το ίδιο ίχνος, έχουμε ότι $\text{tr}(A) = \lambda$ και προκύπτει η πρώτη ισότητα σχετικά με το χαρακτηριστικό πολυώνυμο.

Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του πίνακα $B = (b_{ij})$ με $b_{ij} = b$ είναι ίσο με $\text{Ch}_B(x) = (-x)^{n-1}(nb - x)$. Επίσης παρατηρούμε ότι $A = B - (b-a)\mathbb{I}_n$. Συνεπώς:

$$\det(A - x\mathbb{I}_n) = \det(-(b-a)\mathbb{I}_n - x\mathbb{I}_n).$$

Αν θέσουμε $y = b - a + x$ η παραπάνω παράσταση είναι ίση με

$$\text{Ch}_A(x) = (-y)^{n-1}(nb - y) = (a - b - x)^{n-1}(nb - b + a - x).$$

Λύση της άσκησης VI.14

Αφού $(p_L(x), p_G(x)) = 1$ υπάρχουν πολυώνυμα $f(x), g(x) \in \mathbb{F}[x]$ ώστε

$$f(x)p_L(x) + g(x)p_G(x) = 1 \text{ συνεπώς } f(G)p_L(G) = \text{Id}_V,$$

από όπου προκύπτει ότι ο $f(G)$ είναι ο αντίστροφος της $p_L(G)$.

Υποθέτουμε τώρα ότι υπάρχει μη μηδενικό $v \in V$ ώστε $Gv = 0$. Γράφουμε

$$p_L(x) = a_0 + a_1x + \dots + a_nx^n$$

και έχουμε ότι αφού $p_L(G)$ είναι ισομορφισμός

$$0 \neq p_L(G)v = a_0v = a_0v + a_1Gv + a_2G^2v + \dots + a_nG^nv.$$

Συνεπώς $a_0 \neq 0$. Αν $w \in \text{Ker } L$, τότε αφού $p_L(L) = 0$ έχουμε

$$a_0w = -(a_1L + a_2L^2 + \dots + a_nL^n)w = 0$$

και αφού $a_0 \neq 0$ έχουμε ότι $w = 0$. Συνεπώς η L είναι 1-1 $L: V \rightarrow V$, άρα είναι και επί.

Λύση της άσκησης VI.15

Ένας τριγωνοποιήσιμος πίνακας στο $\mathbb{R}$ έχει όλα τα στοιχεία του στο $\mathbb{R}$ άρα και τα διαγώνια στοιχεία που είναι πραγματικοί αριθμοί. Ένας πίνακας αποτυγχάνει να είναι τριγωνοποιήσιμος στο $\mathbb{R}$ αν το χαρακτηριστικό του πολυώνυμο δεν έχει όλες τις ρίζες πραγματικές. Συνεπώς έχει μια πραγματική ρίζα και δύο συζυγείς μιγαδικές στο $\mathbb{C}$. Άρα στους μιγαδικούς αριθμούς έχει τρεις διαφορετικές ρίζες και είναι διαγωνοποιήσιμος.

Λύση της άσκησης VI.16

Θεωρούμε την βάση του $\mathbb{R}_2[x]$ αποτελούμενη από τα πολυώνυμα $1, x^2 + x + 1, x$.

Παρατηρούμε ότι

$$f(x) = ax^2 + bx = a(x^2 + x + 1) + (b-a)x - a.$$

Άρα ο πίνακας της γραμμικής συνάρτησης ως προς την παραπάνω διατεταγμένη βάση είναι ο

$$\begin{pmatrix} 0 & 0 & -a \\ 0 & 1 & a \\ 0 & 0 & b-a \end{pmatrix}$$

ο οποίος είναι άνω τριγωνικός. Άρα η απεικόνιση είναι τριγωνίσιμη.

Εναλλακτικά μπορούσαμε να διαλέξουμε μια άλλη βάση του χώρου, να γράψετε τον πίνακα και να υπολογίσουμε ότι οι ιδιοτιμές είναι $\{1, 0, b-a\}$ που όλες είναι πραγματικές άρα ο πίνακας είναι τριγωνίσιμος.

Αφού το $a \neq b$ το $b-a \neq 0$. Άρα αν $b-a \neq 1$ τότε ο πίνακας έχει 3 διαφορετικές ιδιοτιμές και είναι διαγωνίσιμος. Στην περίπτωση $b-a=1$ τότε το χαρακτηριστικό πολυώνυμο είναι το $-x(x-1)^2$. Το ελάχιστο πολυώνυμο λοιπόν είναι ή το $x(x-1)$ και ο πίνακας είναι διαγωνίσιμος ή το $x(x-1)^2$ και ο πίνακας δεν είναι διαγωνίσιμος.

Υπολογίζουμε ότι

$$A(A - \mathbb{I}_3) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}$$

από όπου προκύπτει ότι ο πίνακας είναι διαγωνίσιμος σε αυτή την περίπτωση ($b-a=1$) αν και μόνο αν $a=0$.

Λύση της άσκησης VI.17

Από την μορφή του ελαχίστου πολυωνύμου έχουμε ότι ο χώρος $\mathbb{F}^6$ διασπάται σε ένα ευθύ άθροισμα δύο αναλλοίωτων υποχώρων $\mathbb{F}^6 = W_1 \oplus W_2$, όπου ο L_{W_1} έχει ελάχιστο πολυώνυμο $(x-2)$ αναγκαστικά χαρακτηριστικό $(x-2)^2$, ενώ ο L_{W_2} έχει ελάχιστο πολυώνυμο $(x-1)^2$ και αναγκαστικά χαρακτηριστικό $(x-1)^4$.

Η πιθανή μορφή Jordan είναι ένας μπλοκ διαγώνιος πίνακας της μορφής

$$\left(\begin{array}{c|c} A_1 & \mathbf{0}_{2,4} \\ \hline \mathbf{0}_{4,2} & A_2 \end{array} \right)$$

όπου οι πίνακες $A_1 \in \mathbb{F}^{2,2}$ και $A_2 \in \mathbb{F}^{4,4}$. Το ελάχιστο πολυώνυμο του A_1 είναι $(x-2)$ άρα ο πίνακας A_1 είναι διαγωνοποιήσιμος της μορφής $\text{diag}(2, 2)$. Από την άλλη το ελάχιστο πολυώνυμο του πίνακα A_2 είναι $(x-1)^2$ άρα ο πίνακας A_2 δεν μπορεί να είναι διαγωνοποιήσιμος. Επίσης δεν μπορεί να έχει ως μπλοκ το $J(1, 3)$ ούτε το $J(1, 4)$ γιατί αυτά έχουν ελάχιστα πολυώνυμα $(x-1)^3$ και $(x-1)^4$ αντίστοιχα. Συνεπώς οι δυνατές μορφές για τον πίνακα είναι οι παρακάτω:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{ή} \quad \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Λύση της άσκησης VI.18

Ο παραπάνω πίνακας είναι μπλοκ διαγώνιος. Θα μελετήσουμε τα κομμάτια ξεχωριστά. Ο πίνακας

$$A_1 = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}$$

έχει χαρακτηριστικό πολυώνυμο $(x-2)^2$. Αν δεν μας ενδιαφέρει να υπολογίσουμε τον πίνακα μετατροπής Q μπορούμε να επιχειρηματολογήσουμε ως εξής: Είτε ο πίνακας είναι διαγώνιος με ιδιοτιμή 2, οπότε το ελάχιστο του είναι το $x-2$. Όμως $A_1 - 2\mathbb{I}_2 \neq \mathbf{0}$ άρα ο πίνακας όμοιος με τον

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Ο πίνακας

$$A_2 = \begin{pmatrix} -2 & 2 & -3 \\ 1 & 3 & 1 \\ 8 & -3 & 8 \end{pmatrix}$$

έχει χαρακτηριστικό πολυώνυμο $(3 - x)^3$. Αν δεν μας ενδιαφέρει και πάλι ο υπολογισμός του πίνακα μετατροπής υπολογίζουμε:

$$B - 3I_3 = \begin{pmatrix} -5 & 2 & -3 \\ 1 & 0 & 1 \\ 8 & -3 & 5 \end{pmatrix}, \quad (B - 3I_2)^2 = \begin{pmatrix} 3 & -1 & 2 \\ 3 & -1 & 2 \\ -3 & 1 & -2 \end{pmatrix}, \quad (B - 3I_3)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Το συμπέρασμα είναι ότι το ελάχιστο πολυώνυμο είναι το $(x - 3)^3$ και ότι ο αρχικός πίνακας είναι όμοιος με τον

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

Λύση της άσκησης VI.19

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο του πίνακα το οποίο είναι $x^2 - 15x + 50 = (x - 5)(x - 10)$. Αφού το χαρακτηριστικό πολυώνυμο έχει δύο διαφορετικές ρίζες ο πίνακας A είναι διαγωνοποιήσιμος.

Παρατηρούμε ότι

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix}$$

από όπου προκύπτει ότι

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = A^n \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Θέλουμε να βρούμε ένα πίνακα Q ώστε $Q^{-1}AQ = \text{diag}(5, 10)$. Για αυτό λύνουμε το ισοδύναμο σύστημα

$$AQ = Q\text{diag}(5, 10) \text{ ισοδύναμα } \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix},$$

από όπου προκύπτει

$$Q = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \text{ και στην συνέχεια } Q^{-1} = \frac{1}{5} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

Το ζητούμενο αποτέλεσμα είναι το

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = Q \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} Q^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5^{n+1}(2^n - 6) \\ 5^{n-1}(2^{n+1} + 3) \end{pmatrix},$$

Λύση της άσκησης VII.1

Ο χώρος των $n \times n$ πινάκων μπορεί να ταυτιστεί με το $\mathbb{C}^{n^2}$. Παρατηρούμε ότι

$$\text{tr}(AB^*) = \sum_{i=1}^n \sum_{v=1}^n a_{iv} \bar{b}_{iv},$$

δηλαδή ταυτίζεται με το ερμητιανό εσωτερικό γινόμενο στον $\mathbb{C}^{n^2}$.

Λύση της άσκησης VII.2

Αν $\langle \cdot, \cdot \rangle$ είναι το κανονικό ερμητιανό εσωτερικό γινόμενο τότε παρατηρούμε ότι

$$\langle x, y \rangle_A = \langle Ax, Ay \rangle_H,$$

οπότε οι ζητούμενες ιδιότητες του εσωτερικού γινομένου ελέγχονται εύκολα με την βοήθεια του ερμητιανού εσωτερικού γινομένου.

Λύση της άσκησης VII.3

Θωρούμε δύο στοιχεία x, y τα οποία τα εκφράζουμε με μοναδικό τρόπο ως γραμμικούς συνδυασμούς

$$x = \sum_{i=1}^n x_i v_i, \quad y = \sum_{j=1}^n y_j v_j$$

και υπολογίζουμε

$$\langle x, y \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i \overline{y_j} \langle v_i, v_j \rangle$$

από όπου προκύπτει το ζητούμενο.

Λύση της άσκησης VII.4

Ορίζουμε ένα εσωτερικό γινόμενο από την εξίσωση $\langle x, y \rangle = y^* Q x$. Η απαίτηση $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ προκύπτει από την επιμεριστική ιδιότητα του πολλαπλασιασμού ως προς την πρόσθεση. Το ίδιο και η ιδιότητα $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$.

Υπολογίζουμε ότι

$$\langle x, y \rangle = y^* Q x, \quad \langle y, x \rangle = x^* Q y,$$

και

$$\overline{\langle y, x \rangle} = \langle y, x \rangle^* = (x^* Q y)^* = y^* Q^* x.$$

Παρατηρούμε ότι η τελευταία ποσότητα είναι ίση με $\langle x, y \rangle = y^* Q x$ αν και μόνο αν $Q^* = Q$. Πράγματι η ποσότητα $e_i^* Q e_j = q_{ij}$.

Τέλος η συνθήκη $x^* Q x > 0$ προκύπτει από την συνθήκη $\langle x, x \rangle > 0$ και $\langle x, x \rangle = 0$ αν και μόνο αν $x = 0$.

Λύση της άσκησης VII.5

Γνωρίζουμε ότι στον χώρο $C[0, 1]$ των συνεχών συναρτήσεων στο διάστημα $[0, 1]$ το

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt$$

είναι ένα εσωτερικό γινόμενο. Αν περιοριστούμε στον υπόχωρο των πολυώνυμων και ιδιαίτερα στην βάση $\{1, x, x^2, \dots\}$ έχουμε ότι

$$\langle x^i, x^j \rangle = \int_0^1 t^{i+j} dt = \frac{1}{i+j+1},$$

από όπου προκύπτει το ζητούμενο.

Λύση της άσκησης VII.6

Θεωρούμε τον χώρο $W = \langle u_1, \dots, u_n \rangle$. Γράφουμε το $v \in V$ ως

$$v = \sum_{v=1}^n \frac{\langle v, u_v \rangle}{\|u_v\|} u_v + w'$$

με $w' \in W^\perp$. Αφού για $w = \sum_{v=1}^n \frac{\langle v, u_v \rangle}{\|u_v\|} u_v$ έχουμε ότι $\langle w, w' \rangle = 0$ από το πυθαγόρειο θεώρημα έχουμε

$$\|v\|^2 = \|w\|^2 + \|w'\|^2 \geq \sum_{v=1}^n \frac{|\langle v, u_v \rangle|^2}{\|u_v\|^2}.$$

Η ισότητα ισχύει αν και μόνο αν $w' = 0$.

Λύση της άσκησης VII.7

Ξεκινάμε την διαδικασία Gram-Schmidt. Θέτουμε $u_1 = 1$ και στην συνέχεια υπολογίζουμε

$$u_2 = x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} 1 = x - \frac{1}{2}.$$

Στην συνέχεια θέτουμε

$$u_3 = x^2 - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle} - \frac{\langle x^2, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 = x^3 - 1/3 - u_2$$

και

$$\begin{aligned} u_4 &= x^3 - \frac{\langle x^3, 1 \rangle}{\langle 1, 1 \rangle} - \frac{\langle x^3, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 - \frac{\langle x^3, u_3 \rangle}{\langle u_3, u_3 \rangle} u_3 \\ &= x^3 - 1/4 - 9/10 u_2 - 3/2 u_3. \end{aligned}$$

Στην συνέχεια υπολογίζουμε ότι

$$\|u_1\| = 1, \|u_2\| = \frac{1}{2\sqrt{3}}, \|u_3\| = \frac{1}{6\sqrt{5}}, \|u_4\| = \frac{1}{20\sqrt{7}},$$

οπότε έχουμε την ζητούμενη ορθοκανονική βάση

$$e_1 = 1, e_2 = 2\sqrt{3}u_2, e_3 = 6\sqrt{5}u_3, e_4 = 20\sqrt{7}u_4.$$

Λύση της άσκησης VII.8

Παρατηρούμε ότι οι πίνακες E_{ij} οι οποίοι έχουν μια μονάδα στην i -γραμμή και στην j -στήλη και σε όλες τις άλλες θέσεις του πίνακα έχουν μηδενικά, αποτελούν μια ορθοκανονική βάση ως προς το εσωτερικό γινόμενο που μόλις ορίσαμε. Από την άλλη οι διαγώνιοι πίνακες είναι ο χώρος που παράγουν οι πίνακες $E_{11}, E_{22}, \dots, E_{nn}$. Ένας πίνακας $A = (a_{ij})$ γράφεται ως

$$A = \sum_{i=1}^n \sum_{j=1}^n a_{ij} E_{ij}$$

ενώ

$$\langle A, E_{ii} \rangle = a_{ii}$$

Συνεπώς το ορθογώνιο συμπλήρωμα του χώρου των διαγωνίων πινάκων αποτελείται από τον διανυσματικό χώρο των πινάκων που έχουν $a_{11} = a_{22} = \dots = a_{nn} = 0$.

Λύση της άσκησης VII.9

Προφανώς οι ιδιοτιμές είναι οι 1, 0, 3. Για κάθε ιδιοτιμή υπολογίζουμε ένα ιδιοδιάνυσμα. Για την ιδιοτιμή 1 ο ιδιόχωρος είναι μονοδιάστατος και ένα ιδιοδιάνυσμα που τον παράγει είναι το $(0, -1, 1)^t$, για την ιδιοτιμή 0 ο ιδιόχωρος είναι επίσης μονοδιάστατος και ένα ιδιοδιάνυσμα που τον παράγει είναι το $(1, 1, 1)^t$ ενώ για την ιδιοτιμή 3 ο υπόχωρος είναι επίσης μονοδιάστατος και ένα ιδιοδιάνυσμα που τον παράγει είναι το $(-2, 1, 1)^t$.

Παρατηρούμε ότι ο πίνακας μας είναι συμμετρικός (πραγματικός ερμητιανός) και συνεπώς οι ιδιοχώροι είναι ανά δύο κάθετοι. Άρα για να βρούμε μια ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα αρκεί να διαιρέσουμε τα διανύσματα που βρήκαμε παραπάνω με τα μέτρα τους. Οπότε μια ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα είναι η

$$\left(\begin{pmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \right)$$

Προσοχή: Δεν θα μπορούσαμε να εφαρμόσουμε την μέθοδο Gram-Schmidt χωρίς με κάποιο τρόπο να δικαιολογήσουμε ότι τα διανύσματα που προκύπτουν είναι ιδιοδιανύσματα. Παρατηρούμε ότι το ιδιοδιάνυσμα $(1, 1, 1)^t$ ανήκει στον πυρήνα του πίνακα, άρα

$$A \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 0.$$

Αφού και οι δύο πίνακες είναι συμμετρικοί αν θεωρήσουμε τον πολλαπλασιασμό με αντίθετη σειρά θα πάρουμε πάλι 0.

Εναλλακτικά θα μπορούσαμε να επιχειρηματολογήσετε και ως εξής: Ο πίνακας έχει χαρακτηριστικό πολυώνυμο ίσο με το ελάχιστο $m_A(x)$, άρα αφού $m_A(A) = 0$ έχουμε ότι $AG = 0$ αν θέσουμε $G = (A - I_3)(A - 3I_3)$. Ο πίνακας G είναι μη μηδενικός γιατί διαφορετικά θα ικανοποιούσε ένα πολυώνυμο βαθμού μικρότερου από το ελάχιστο. Τέλος παρατηρήστε ότι ο πίνακας G αντιμετατίθεται με τον πίνακα A , αφού ο πίνακας G είναι πολυωνυμική έκφραση του A .

Επίσης θα μπορούσαμε να γράψουμε τον πίνακα $A = P \text{diag}(0, 1, 3) P^{-1}$ και να θέσουμε

$$G = P \text{diag}(1, 0, 0) P^{-1}.$$

Τέλος θα μπορούσαμε να θεωρήσουμε τον χώρο που παράγουν οι γραμμές οι ή στήλες (είναι ο ίδιος αφού ο A είναι συμμετρικός) να υπολογίσουμε τον κάθετο υπόχωρο και να πάρουμε έναν συμμετρικό πίνακα που να έχει γραμμές διανύσματα από τον κάθετο υπόχωρο.

Λύση της άσκησης VII.10

Υπολογίζουμε ότι

$$\begin{aligned} \langle L_M(A), B \rangle &= \text{tr} B^* M A = \text{tr} M A B^* = \text{tr} (A B^* M) \\ &= \text{tr} (A (M^* B)^*) = \text{tr} (M^* B)^* A = \langle A, M^* B \rangle = \langle A, L_{M^*} B \rangle. \end{aligned}$$

Λύση της άσκησης VII.11

Υπολογίζουμε ότι

$$\begin{aligned} \langle M_f g, h \rangle &= \langle fg, h \rangle = \int_0^1 f(t) g(t) \overline{h(t)} dt \\ &= \int_0^1 g(t) \overline{f(t)} h(t) dt = \langle g, \bar{f} h \rangle = \langle g, M_{\bar{f}}(h) \rangle. \end{aligned}$$

Για την συνάρτηση παραγώγισης D παρατηρούμε ότι

$$\langle Df, g \rangle = f(1)g(1) - f(0)g(0) - \langle f, Dg \rangle.$$

Για g σταθερό υποθέτουμε ότι υπάρχει πολυώνυμο D^* ώστε $\langle Df, g \rangle = \langle f, D^*g \rangle$. Τότε

$$\langle f, D^*g \rangle = f(1)g(1) - f(0)g(0) - \langle f, Dg \rangle$$

ισοδύναμα

$$\langle f, D^*g + Dg \rangle = f(1)g(1) - f(0)g(0).$$

Παρατηρούμε ότι για g σταθερό το $L(f) = f(1)g(1) - f(0)g(0)$ είναι ένα στοιχείο του δυϊκού χώρου όπως στην παρατήρηση **VII.7.4** και δεν υπάρχει $h \in \mathbb{C}[x]$ ώστε $L(f) = \langle f, h \rangle$ εκτός αν $L = 0$. Συνεπώς η ύπαρξη του D^*g μας δίνει ότι το $h = D^*g + Dg$ πρέπει να είναι μηδενικό και $g(0) = g(1) = 0$. Αν διαλέξουμε g ώστε $g(0) \neq 0$ ή $g(1) \neq 0$ το D^*g δεν ορίζεται και δεν υπάρχει συζυγής.

Λύση της άσκησης VII.12

Είναι σαφές ότι το γινόμενο δύο κάτω τριγωνικών πινάκων είναι κάτω τριγωνικός πίνακας αλλά και ότι το γινόμενο δύο πινάκων που ο καθένας έχει θετικά στοιχεία στην διαγώνιο θα έχει και αυτό θετικά στοιχεία στην διαγώνιο.

Θεωρούμε ένα κάτω τριγωνικό αντιστρέψιμο πίνακα B τον οποίο τον γράφουμε στην μορφή

$$B = \Delta(\mathbb{I}_n + L)$$

όπου ο Δ είναι διαγώνιος πίνακας με μή-μηδενικά στοιχεία ενώ ο L είναι κάτω τριγωνικός με μηδενικά στοιχεία στην διαγώνιο. Ισχύει ότι $L^n = 0$ (το οποίο το βλέπουμε είτε από την διαγώνια μορφή του πίνακα είτε εξυπνότερα από το Θεώρημα Caley-Hamilton). Τώρα υπολογίζουμε ότι

$$(\mathbb{I}_n + L)(\mathbb{I}_n - L + L^2 - L^3 + \dots + (-1)^{n-1}L^{n-1}) = \mathbb{I}_n.$$

Άρα

$$B^{-1} = (\mathbb{I}_n + L)^{-1}\Delta^{-1} = (\mathbb{I}_n - L + L^2 - L^3 + \dots + (-1)^{n-1}L^{n-1})\Delta^{-1}$$

ο οποίος είναι κάτω τριγωνικός, ως άθροισμα γινομένων κάτω τριγωνικών πινάκων. Η απαίτηση για θετικά στοιχεία στην διαγώνιο είναι εμφανής από τον υπολογισμό αντιστρόφου διαγώνιου πίνακα.

Λύση της άσκησης VII.13

Το ότι το $\langle \cdot, \cdot \rangle_2$ είναι εσωτερικό γινόμενο είναι μια εύκολη επαλήθευση.

Παρατηρούμε ότι

$$\langle f, g \rangle_2 = \int_0^1 f(t)g(t)t^2 dt = \langle Tf, Tg \rangle_1.$$

Η συνάρτηση $: W \rightarrow V$ είναι 1-1 αλλά επειδή η διάστασις των V, W δεν είναι πεπερασμένες δεν είναι επί. Πράγματι δεν υπάρχει τρόπος να πάρουμε την σταθερή συνάρτηση $1 : [0, 1] \rightarrow \mathbb{R}$ ως γινόμενο $χg(x) = 1$, η συνάρτηση $g(x) = 1/x$ δεν ορίζεται στο 0.

Λύση της άσκησης VII.14

Ένας 1×1 πίνακας (a) είναι ορθογώνιος αν και μόνο αν $a^2 = 1$ δηλαδή αν και μόνο αν $a = \pm 1$. Ένας 1×1 πίνακας (a) είναι μοναδιαίος αν και μόνο αν $a\bar{a} = \|a\|^2 = 1$ δηλαδή το a είναι ένας μιγαδικός πάνω στον μοναδιαίο κύκλο, $a = e^{i\theta}$, $\theta \in \mathbb{R}$.

Έστω τώρα ένας 2×2 πίνακας

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ με } A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Αν ο πίνακας είναι ορθογώνιος, δηλαδή αν $AA^t = \mathbb{I}_n$, τότε $\det(A) = \pm 1$, αφού $\det A^t = \det A$. Συνεπώς ο πίνακας είναι ορθογώνιος αν

$$\det A = -1 \text{ και } \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -d & c \\ b & -a \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

$$\det A = 1 \text{ και } \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -c \\ b & a \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

και στις δύο περιπτώσεις ο υπολογισμός της ορίζουσας δίνει την σχέση $a^2 + b^2 = 1$.

Στην περίπτωση ενός πραγματικού πίνακα έχουμε ότι αυτός είναι ένας πίνακας της μορφής

$$A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Στην μιγαδική περίπτωση τώρα, παρατηρούμε ότι για μοναδιαίους πίνακες η σχέση $AA^* = I_2$, δίνει ότι $\det A \det A = 1$ συνεπώς η ορίζουσα έχει μέτρο 1, δηλαδή $\det A = e^{i\theta}$, $\theta \in \mathbb{R}$. Επιπλέον ο πίνακας είναι μοναδιαίος αν και μόνο αν

$$\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} -d & -b \\ -c & -a \end{pmatrix} \Rightarrow = \begin{pmatrix} a & b \\ -e^{i\theta}\bar{b} & e^{i\theta}\bar{a} \end{pmatrix}$$

και $a^2 + |b|^2 = 1$.

Λύση της άσκησης VII.15

1. Παρατηρούμε ότι

$$\text{ch}_{A^*}(x) = \det(A^* - xI_n) = \overline{\det(A^t - \bar{x}I_n)} = \overline{\det(A - \bar{x}I_n)} = \overline{\text{ch}_A(\bar{x})} = \overline{\text{ch}_A}(x).$$

2. Έχουμε ότι αν $f(x) = \sum_{v=0}^n a_v x^v \in \mathbb{C}[x]$ πολυώνυμο ώστε $f(A) = 0$ τότε

$$0 = 0^* = \sum_{v=0}^n a_v (A^t)^v = \sum_{v=0}^n \bar{a}_v (A^*)^v.$$

Συνεπώς αν $m_A(x)$ είναι το ελάχιστου βαθμού πολυώνυμο που μηδενίζει τον A , το $\overline{m}_A(x)$ είναι το ελάχιστου βαθμού πολυώνυμο που μηδενίζει το A^* .

3. Σαφές από το [II.1](#).

Λύση της άσκησης VII.16

Παρατηρούμε ότι αν $B = (b_{ij})$ τότε $b_{ij} = x_i x_j$. Υπολογίζουμε ότι αν $B^2 = (c_{ij})$ τότε

$$c_{ij} = \sum_{v=1}^n b_{iv} b_{vj} = \sum_{v=1}^n x_i x_v^2 x_j = b_{ij} \|x\|^2.$$

Αφού $B^2 = \|x\|B$ έχουμε ότι για το πολυώνυμο

$$f(x) = x^2 - (2 + \|x\|^2)x + 1 + \|x\| = (x-1)(x - (1 + \|x\|^2))$$

ισχύει ότι $f(A) = 0$. Επειδή δε κανένας από τους δύο γραμμικούς παράγοντες δεν μηδενίζουν τον A αυτό είναι το ελάχιστο.

Για τον ιδιόχωρο της ιδιοτιμής 1 έχουμε ότι αποτελείται από τα διανύσματα v ώστε

$$(A - I_n)v = x^t x v = 0$$

και ότι κάθε διάνυσμα v που είναι κάθετο στο x ανήκει στον ιδιόχωρο αυτό. Άρα ο ιδιοχώρος της ιδιοτιμής 1 έχει διάσταση $\geq n-1 = \dim\langle x \rangle^\perp$. Αφού ο ιδιόχωρος της ιδιοτιμής $1 + \|x\|^2$ έχει διάσταση ≥ 1 έχουμε ότι οι ιδιοστάσεις των ιδιοχώρων είναι $n-1$ και 1 αντίστοιχα. Τέλος η ορίζουσα είναι το γινόμενο των ιδιοτιμών άρα ίση με $1 + \|x\|^2$.

Λύση της άσκησης VII.17

Αν A δεν είναι αντιστρέψιμος η ανισότητα ισχύει με προφανή τρόπο και γίνεται ισότητα στην περίπτωση που κάποιο $a_i = 0$.

Αν ο A είναι αντιστρέψιμος τότε από την πρόταση VII.8.12 υπάρχει μοναδικός κάτω τριγωνικός πίνακας B με θετικά στοιχεία στην διαγώνιο ώστε ο πίνακας $BA = U$ να είναι μοναδιαίος. Έχουμε ότι

$$|\det U|^2 = \det U \det U^* = \det UU^* = \det I_n = 1.$$

Γράφουμε

$$|\det(A)| = |\det(B^{-1}U)| = |\det(B^{-1})| = b_{11}^{-1} \cdots b_{nn}^{-1}$$

από την άλλη έχουμε ότι πίνακας $B^{-1} = (c_{ij})$ είναι κάτω τριγωνικός με $c_{ii} = b_{ii}^{-1}$.

$$\begin{aligned} a_i &= (c_{i1}, c_{i2}, \dots, b_{ii}^{-1}, 0, \dots, 0)U \\ &= c_{i1}U^1 + c_{i2}U^2 + \dots + b_{ii}^{-1}U^i \text{ όπου } U^1, \dots, U^n \text{ οι στήλες του } U. \end{aligned}$$

Αρα

$$\|a_i\|^2 = |c_{i1}|^2 \|U^1\|^2 + \dots + |b_{ii}^{-1}|^2 \|U^i\|^2 = |c_{i1}|^2 + \dots + |b_{ii}^{-1}|^2 \geq |b_{ii}^{-1}|^2.$$

Από την τελευταία σχέση προκύπτει ότι

$$|a_1| \cdots |a_n| \geq b_{11}^{-1} \cdots b_{nn}^{-1} = |\det(A)|.$$

Τέλος, παρατηρούμε ότι η προηγούμενη ανισότητα γίνεται ισότητα αν και μόνο αν $c_{i1} = \dots = c_{i,i-1} = 0$, ισοδύναμα αν και μόνο αν $a_i = b_{ii}^{-1}U^i$, αν και μόνο αν οι γραμμές a_i είναι ανά δύο ορθογώνιες.

Λύση της άσκησης VII.18

Άμεση συνέπεια της άσκησης VII.17.

Λύση της άσκησης VII.19

- Υπολογίζουμε ότι ο πίνακας A έχει ιδιοτιμές 1, 4. Ένα ιδιοδιάνυσμα της ιδιοτιμής 4 είναι το $(1, 1)^t$. Συμπληρώνουμε σε βάση του $\mathbb{R}^2$ με το $\{(1, 1)^t, (1, 0)^t\}$. Στην συνέχεια ορθοκανονικοποιούμε στην ορθοκανονική βάση

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

και σχηματίζουμε τον μοναδιαίο πίνακα:

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

Παρατηρούμε ότι

$$U^*AU = \begin{pmatrix} 4 & -1 \\ 0 & 1 \end{pmatrix}.$$

Προσοχή: Ο πίνακας είναι διαγωνίσιμος αφού έχει δύο διαφορετικές ιδιοτιμές. Όμως δεν είναι δυνατόν να έρθει σε διαγώνια μορφή μέσω μοναδιαίου πίνακα, αφού δεν είναι κανονικός. Αν απαιτήσουμε να κάνουμε αλλαγή βάσης μέσω μοναδιαίου πίνακα μπορούμε να τον φέρουμε μόνο σε άνω τριγωνική μορφή.

Εναλλακτικά θα μπορούσαμε να δουλέψουμε με την ιδιοτιμή 1 και να συμπληρώσουμε ένα ιδιοδιάνυσμά της σε ορθοκανονική βάση του $\mathbb{R}^2$.

2. Κάνοντας χρήση του φασματικού θεωρήματος για το πολυώνυμο $f(x) = x^{15} - 3x^9 + 15$, βλέπουμε ότι ο ζητούμενος πίνακας έχει ιδιοτιμές $f(1)$, $f(4)$ και τα ίδια ιδιοδιανύσματα.

Λύση της άσκησης VII.20

1. Γνωρίζουμε ότι ο πίνακας Q θα είναι της μορφής:

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & \mathbb{U} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Πράγματι, παρατηρούμε ότι

$$Q^{-1}BQ = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{3}{\sqrt{2}} \\ 0 & 4 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του B είναι το $-x^3 + 6x^2 - 9x + 4$ το οποίο από το θεώρημα Caley-Hamilton θα πρέπει να μηδενίζεται από τον B , δηλαδή

$$-B^3 + 6B^2 - 9B + 4\mathbb{I}_3 = 0,$$

οπότε πολλαπλασιάζοντας με B^{-1} καταλήγουμε στο

$$B^{-1} = \frac{1}{4}(B^2 - 6B + 9\mathbb{I}_3).$$

Λύση της άσκησης VII.21

Παρατηρούμε ότι ο πίνακας $\Gamma + \Gamma^*$ είναι ερμητιανός άρα έχει πραγματικές ιδιοτιμές. Αν ο ζητούμενος πίνακας δεν ήταν αντιστρέψιμος αυτό θα σήμαινε ότι

$$\det(\Gamma + \Gamma^* + i\mathbb{I}_n) = 0$$

το οποίο θα σήμαινε ότι το $-i$ θα ήταν μια ρίζα του χαρακτηριστικού πολυωνύμου, άτοπο.

Λύση της άσκησης VIII.2

Έστω $L : V \rightarrow V$ η γραμμική συνάρτηση ώστε $f(v, w) = \langle Lv, w \rangle$. Σύμφωνα με την πρόταση [VII.9.11](#) υπάρχει ορθοκανονική βάση $B = \{e_1, \dots, e_n\}$ στην οποία ο (L, B, B) να είναι άνω τριγωνικός. Έχουμε ότι

$$f(e_i, e_j) = \langle Le_i, e_j \rangle = 0,$$

για $j > i$.

Λύση της άσκησης VIII.3

Αν η f είναι ερμητιανή τότε $f(v, v) = \overline{f(v, v)}$ για κάθε $v \in V$, συνεπώς $f(v, v) \in \mathbb{R}$.

Αντιστρόφως, έστω $v, w \in V$. Θα πρέπει να δείξουμε ότι $f(v, w) = \overline{f(w, v)}$ για κάθε $v, w \in V$. Έχουμε ότι

$$f(v, w) + f(w, v) = f(v + w, v + w) - f(v, v) - f(w, w) \in \mathbb{R}.$$

Επίσης

$$-if(v, w) + if(w, v) = f(v + iw, v + iw) - f(v, v) - f(w, w) \in \mathbb{R}.$$

Συνεπώς

$$f(v, w) + f(w, v) = \overline{f(v, w)} + \overline{f(w, v)}, \quad -if(v, w) + if(v, w) = \overline{if(v, w)} - \overline{if(w, v)}.$$

Αν πολλαπλασιάσουμε την δεύτερη εξίσωση με i και την προσθέσουμε στην πρώτη έχουμε ότι

$$2f(v, w) = 2f(w, v).$$

Λύση της άσκησης VIII.4

Υποθέτουμε ότι η L έχει την φασματική ανάλυση

$$L = \lambda_1 \pi_1 + \cdots + \lambda_k \pi_k.$$

Η L είναι αυτοσυζυγής αν και μόνο αν $L = L^*$ οπότε η παραπάνω εξίσωση μας δίνει:

$$(\lambda_1 - \bar{\lambda}_1) \pi_1 + \cdots + (\lambda_k - \bar{\lambda}_k) \pi_k = 0.$$

Η παραπάνω σχέση είναι αληθής αν και μόνο αν $\lambda_i = \bar{\lambda}_i$ για κάθε i . Πράγματι αν $\lambda_i = \bar{\lambda}_i$ τότε η σχέση είναι αληθής και αν η σχέση είναι αληθής τότε συνθέτοντας με την π_j και χρησιμοποιώντας ότι καμία π_j δεν είναι η μηδενική συνάρτηση έχουμε $\lambda_i = \bar{\lambda}_i$.

Για να εξετάσουμε τις μη-αρνητικές κανονικές γραμμικές συναρτήσεις υπολογίζουμε

$$\langle Lv, v \rangle = \left\langle \sum_{j=1}^k \lambda_j \pi_j v, \sum_{j=1}^k \pi_j v \right\rangle = \sum_{j=1}^k \sum_{i=1}^k \lambda_j \langle \pi_j v, \pi_i v \rangle = \sum_{j=1}^k \lambda_j \|\pi_j v\|^2.$$

Στην παραπάνω σχέση χρησιμοποιήσαμε το ότι $\langle \pi_j v, \pi_i v \rangle = 0$ για $i \neq j$. Συνεπώς η σχέση $\langle Lv, v \rangle \geq 0$ ισχύει αν και μόνο αν $\lambda_j \geq 0$ για κάθε j .

Τέλος για τις μοναδιαίες γραμμικές συναρτήσεις έχουμε ότι

$$LL^* = |\lambda_1|^2 \pi_1 + \cdots + |\lambda_k|^2 \pi_k,$$

οπότε αν $LL^* = \text{Id}_V$ τότε

$$\text{Id}_V = |\lambda_1|^2 \pi_1 + \cdots + |\lambda_k|^2 \pi_k,$$

οπότε συνθέτοντας με π_i έχουμε

$$\pi_i = |\lambda_i|^2 \pi_i,$$

και αφού $\pi_i \neq 0$ καταλλήλουμε στο $|\lambda_j| = 1$ για κάθε j .